



**ARQUITECTURA DE INTELIGENCIA DE NEGOCIOS PARA INTEGRAR
RESULTADOS DE LA ESTRATEGIA “EVALUAR PARA AVANZAR” Y LAS
CONDICIONES INSTITUCIONALES DE PLANTELES EDUCATIVOS OFICIALES DEL
MUNICIPIO DE MONTERÍA**

Daniel Andrés Alarcon Solera

Universidad Ean

Facultad de ingenierías

Maestría en inteligencia de negocios

Ciudad, Colombia

17/06/2025

ARQUITECTURA DE INTELIGENCIA DE NEGOCIOS PARA INTEGRAR
RESULTADOS DE LA ESTRATEGIA “EVALUAR PARA AVANZAR” Y
LAS CONDICIONES INSTITUCIONALES DE PLANTELES
EDUCATIVOS OFICIALES DEL MUNICIPIO DE MONTERÍA

2

**ARQUITECTURA DE INTELIGENCIA DE NEGOCIOS PARA INTEGRAR
RESULTADOS DE LA ESTRATEGIA “EVALUAR PARA AVANZAR” Y LAS
CONDICIONES INSTITUCIONALES DE PLANTELES EDUCATIVOS OFICIALES DEL
MUNICIPIO DE MONTERÍA**

Daniel Andrés Alarcon Solera

Trabajo de grado presentado como requisito para optar al título de

Magister en Inteligencia de Negocios

Director (a):

CAROLINA MARIA LUQUE ZABALA

Modalidad:

Monografía

Universidad Ean

Facultad de ingenierías

Maestría en inteligencia de negocios

Montería, Colombia

17/06/2025

Nota de aceptación:

Firma del jurado

Firma del jurado

Firma del director del trabajo de grado

Ciudad, día/mes/año

A aquellos que, con su paciencia infinita y amor incondicional, han iluminado mi camino. A mi familia y amigos, quienes han sido mi sostén y fuente de inspiración a lo largo de este viaje.

"Explorar el conocimiento es adentrarse en el universo de lo posible. Reconocer nuestras limitaciones y aceptarlas nos guía hacia la verdadera sabiduría." - Proverbio

Anónimo

Agradecimientos

Quiero expresar mi profundo agradecimiento a aquellas personas cuya colaboración fue fundamental en la realización de este trabajo de grado. Sus aportes generosos y su apoyo incondicional han sido piedras angulares en el desarrollo de este proyecto.

Agradezco sinceramente a Nuvis Ines Berrocal Paternina, magister y profesional especializada en la secretaria de educación de Montería, por su dedicación y colaboración constante. Su orientación y asistencia fueron vitales para el éxito de este proyecto, convirtiéndola en la persona más colaborativa y comprometida en mi camino.

Un agradecimiento especial a Daniela Alarcón Solera, Magíster en Didáctica de Ciencias y docente, quien ha sido un faro de sabiduría y un pilar de apoyo tanto en lo académico como en lo personal. Su astucia y compromiso han enriquecido enormemente este trabajo.

Asimismo, quiero expresar mi gratitud hacia Carolina María Luque Zabala, asesora de la tesis, cuya guía sabia y perspicaz ha sido un faro en este proceso académico. Su compromiso y orientación fueron fundamentales para dar forma y dirección a esta investigación. Su enfoque comprometido y su aguda percepción han sido fundamentales para elevar la calidad de este trabajo, convirtiéndola en una guía excepcional a lo largo de esta travesía académica.

Cada uno de ustedes ha dejado una huella imborrable en este proyecto, y estoy sinceramente agradecido por su dedicación, paciencia y valiosos aportes que han contribuido al éxito de este trabajo de grado.

Resumen

Esta investigación tiene como objetivo proponer una arquitectura de inteligencia de negocios (BI) que integre los resultados de la estrategia “Evaluar para avanzar” y las condiciones institucionales de los centros educativos oficiales de Montería para los años 2021-2022. Donde se busca revelar patrones que contribuyan al fortalecimiento de la toma de decisiones en políticas educativas. Para llevarlo a cabo se establecieron objetivos como la búsqueda de diferentes referentes teóricos, la caracterización de fuentes de datos y la aplicación de procesos de extracción, transformación y carga (ETL) para la estrategia y condiciones institucionales. También se busca configurar un sistema de almacenamiento de datos y generar visualizaciones que consoliden los resultados.

La premisa destaca que el desarrollo de una arquitectura de inteligencia de negocios para la integración de datos se presenta como una estrategia efectiva, capaz de examinar relaciones y patrones, lo cual repercute positivamente en la toma de decisiones en políticas educativas. Además, la investigación ofrece pautas metodológicas para futuros diseños de arquitecturas de BI en el sector educativo colombiano.

En el proceso, se emplearon metodologías y técnicas específicas, destacándose la utilización de programación en Python para la manipulación de datos y análisis relacional. Además, se aprovechó Power BI para la visualización de resultados y análisis exploratorio de datos. Estas estrategias proporcionaron una interpretación significativa de las variables analizadas, respaldando de manera concluyente la hipótesis inicial planteada.

Palabras clave: Inteligencia de negocios, arquitectura de inteligencia de negocios, estrategia “Evaluar para avanzar”, condiciones institucionales, análisis de datos, indicadores claves, indicadores de relación, Python, Power BI, Montería

Abstract

This research aims to propose a business intelligence (BI) architecture that integrates the results of the “Evaluar para avanzar” strategy and the institutional conditions of public schools in Montería for the years 2021–2022. The study seeks to identify patterns that contribute to strengthening decision-making in educational policies. To achieve this, objectives were established, including reviewing relevant theoretical frameworks, characterizing data sources, and applying ETL (Extract, Transform, Load) processes for the assessment strategy and institutional conditions. Additionally, the study aims to configure a data storage system and generate visualizations to consolidate findings.

The premise emphasizes that developing a business intelligence architecture for data integration serves as an effective strategy to examine relationships and patterns, thereby positively impacting educational policy decisions. Furthermore, the research provides methodological guidelines for future BI architecture designs in the Colombian educational sector.

The process employed specific methodologies and techniques, notably Python programming for data manipulation and relational analysis, alongside Power BI for data visualization and exploratory analysis. These strategies enabled a meaningful interpretation of the analyzed variables, conclusively supporting the initial hypothesis.

Keywords: Business Intelligence, Business Intelligence Architecture, “Evaluar para avanzar”, Institutional Conditions, Data Analysis, Key Indicators, Relationship Indicators, Montería.

Contenido

	Pág.
1. Introducción.....	12
2. Objetivos	15
2.1. Objetivo general	15
2.2. Objetivos específicos	15
3. Justificación.....	16
4. Marco Teórico	18
4.1. Inteligencia de negocios.	18
4.2. Arquitectura de inteligencia de negocios.....	18
4.2.1. Capa de fuente de datos.....	21
4.2.2. Capa extracción, transformación y carga (ETL).....	21
4.2.4. Capa de análisis	22
4.2.5. Capa de distribución	23
4.3. Estrategia “Evaluar para avanzar” 3° a 11° (EPA).....	24
4.4. Condiciones institucionales.....	27
4.4.1. Condiciones institucionales en relación con los resultados académicos	28
4.5. Inteligencia de negocios en un marco de aplicación al sector educativo y estrategias de evaluación y seguimiento.....	29
4.5.1. Estudios en el ámbito internacional	30
4.5.2. Estudios en el ámbito nacional	31

ARQUITECTURA DE INTELIGENCIA DE NEGOCIOS PARA INTEGRAR RESULTADOS DE LA ESTRATEGIA “EVALUAR PARA AVANZAR” Y LAS CONDICIONES INSTITUCIONALES DE PLANTELES EDUCATIVOS OFICIALES DEL MUNICIPIO DE MONTERÍA	9
4.5.3. Antecedentes de estudios analíticos de la EPA	33
4.6. Métodos de analítica de datos	34
5. Hipótesis	36
6. Fuente de datos y Variables	37
6.1. Fuente de datos.....	37
6.2. Variables.....	39
7. Diseño metodológico.....	41
7.1. Enfoque de la investigación	42
7.3. Población y muestra	43
7.4. Instrumentos.....	43
7.5. Herramientas tecnológicas para la construcción de la arquitectura de BI.....	44
7.6. Técnicas para el análisis de la información.....	46
7.6.1. Indicadores Claves de Desempeño (KPI)	47
7.6.2. Modelo de regresión	57
8. Diseño de la arquitectura BI	58
8.1. Capa de fuente de datos.....	58
8.2. Capa ETL (Extracción, transformación y carga).....	58
8.2.1. Extracción.....	58
8.2.2. Transformación.....	58
8.2.3. Carga.....	58

ARQUITECTURA DE INTELIGENCIA DE NEGOCIOS PARA INTEGRAR RESULTADOS DE LA ESTRATEGIA “EVALUAR PARA AVANZAR” Y LAS CONDICIONES INSTITUCIONALES DE PLANTELES EDUCATIVOS OFICIALES DEL MUNICIPIO DE MONTERÍA	10
8.3. Capa de Análisis.....	62
8.3.1. Análisis exploratorio de datos (EDA) mediante indicadores claves	63
8.3.2. Análisis del modelo relacional.....	63
8.4. Capa de Distribución	64
8.4.1. Indicadores de desempeño académico.....	65
8.4.2. Indicadores de recursos humanos	66
8.4.3. Indicadores de equidad	67
8.4.4. Indicadores socioeconómicos.....	68
8.4.5. Indicadores de relación entre condiciones institucionales y resultados de la prueba	68
8.4.6. Modelo predictivo de resultados	69
9. Trabajo de Campo.....	70
9.1. Procesamiento de los datos.....	70
9.2. Análisis de indicadores	71
9.2.1. Indicadores de rendimiento.....	71
9.2.2. indicadores de recursos humanos	75
9.2.3. indicador de equidad	81
9.2.4. Indicador socioeconómico	84
9.2.5. Indicadores de relación entre las condiciones institucionales y los resultados de desempeño de la estrategia EPA.	86
9.3. Modelo predictivo.....	99
9.3.1. Propuesta de solución al problema planteado	102

ARQUITECTURA DE INTELIGENCIA DE NEGOCIOS PARA INTEGRAR RESULTADOS DE LA ESTRATEGIA “EVALUAR PARA AVANZAR” Y LAS CONDICIONES INSTITUCIONALES DE PLANTELES EDUCATIVOS OFICIALES DEL MUNICIPIO DE MONTERÍA 11

10. Directrices metodológicas para diseñar arquitecturas de inteligencia de negocios en el sector educativo en Colombia.....	104
11. Conclusiones, recomendaciones y trabajos futuros.....	107
11.1. Conclusiones	107
11.2. Recomendaciones	108
11.3. Trabajos futuros.....	109
12. Referencias	111

Lista de Figuras

	Pág.
Figura 1. Arquitectura de inteligencia de negocios	20
Figura 2. Instrumentos de valoración de competencias básicas y cuestionarios auxiliares	26
Figura 3. Diseño metodológico.....	41
Figura 4. Cuadrantes de relación.	54
Figura 5. Modelo Físico de datos.	60
Figura 6. Arquitectura BI aplicada en esta monografía.....	71
Figura 7. Desempeño académico promedio por rango.....	72
Figura 8. Desempeño académico promedio según zona 2021 (izquierda) y 2022 (derecha).....	72
Figura 9. Desempeño académico promedio según jornada, año 2021.....	73
Figura 10. Brecha de género en los resultados de la prueba EPA por grado.	74
Figura 11. Desempeño promedio de los distintos instrumentos (años conjuntos).	75
Figura 12. Distribución docente por género año 2021 (superior) y 2022 (inferior).	76
Figura 13. Proporción docente según contrato.....	77
Figura 14. Proporción docente según rango etario.....	78
Figura 15. Proporción de docentes por área y desempeño, año 2021.....	79
Figura 16. Proporción de docentes por área y desempeño, año 2022.....	80
Figura 17. Distribución de estudiantes en condición especial.....	82
Figura 18. Distribución de instituciones según estudiantes en condición especial.....	82
Figura 19. Desempeño académico promedio según tipo de condición especial.	84
Figura 20. Proporción de estudiantes según estrato.	85
Figura 21. Desempeño académico promedio según estrato.....	86

ARQUITECTURA DE INTELIGENCIA DE NEGOCIOS PARA INTEGRAR RESULTADOS DE LA ESTRATEGIA “EVALUAR PARA AVANZAR” Y LAS CONDICIONES INSTITUCIONALES DE PLANTELES EDUCATIVOS OFICIALES DEL MUNICIPIO DE MONTERÍA	11
Figura 22. Indicador ICAE (Años conjuntos 2021 y 2022).	87
Figura 23. Indicadores IDAD y IEDR (Años conjuntos 2021 y 2022).	88
Figura 24. Indicadores IDRRA y IRPDD (Años conjuntos 2021 y 2022).	89
Figura 25. Indicadores IRNDP y I RNDS (Años conjuntos 2021 y 2022).	90
Figura 26. Indicador IDED (Años conjuntos 2021 y 2022).	91
Figura 27. Indicadores ICAE e IDED, para el año 2021.	92
Figura 28. Indicadores IEDR e IDAD, para el año 2021.	93
Figura 29. Indicadores ICAE e IDED, para el año 2022.	94
Figura 30. Indicadores IEDR e IDAD, para el año 2022.	95
Figura 31. Comparativo indicador ICAE para la institución educativa Caño Viejo Palotal en los años 2021-2022 (izquierda-derecha).	96
Figura 32. Comparativo indicador IDRRA para la institución educativa Caño Viejo Palotal en los años 2021-2022 (izquierda-derecha).	98
Figura 33. Matriz de correlación año.	99
Figura 34. Condiciones institucionales sensibles dentro del modelo de predicción. .	101
Figura 35. Predicción desempeño promedio del municipio de Montería.	102

Lista de Tablas

	Pág.
Tabla 1. Variables de la estrategia EPA	39
Tabla 2. Variables de las condiciones institucionales	39
Tabla 3. Tecnologías para la construcción de una arquitectura BI.....	44
Tabla 4. Indicadores de desempeño académico.....	48
Tabla 5. Indicadores de recursos humanos.....	50
Tabla 6. Indicadores de equidad educativa.....	51
Tabla 7. Indicadores socioeconómicos.....	52
Tabla 8. Indicadores de relación entre desempeño académico y las condiciones institucionales.....	55
Tabla 9. Descripción modelo conceptual, Hechos y Dimensiones.....	59
Tabla 10. Métricas de rendimiento modelo de predicción.....	100

1. Introducción

La transformación digital ha impactado significativamente todos los sectores, incluyendo el educativo. La implementación de tecnologías de la información y comunicación (TIC) ha llevado no solo a la automatización de procesos, sino también a la generación de volúmenes masivos de datos que requieren gestión eficiente, disponibilidad y actualización constante (Gonzaga, 2021). En este contexto, la inteligencia de negocios (BI, por sus siglas en inglés) ha emergido como una herramienta clave para facilitar la gestión de datos y la toma de decisiones fundamentadas en evidencia (Salluco, 2021). En el ámbito educativo, estudios recientes han demostrado que la BI permite optimizar la gestión académica, mejorar la calidad del aprendizaje y fortalecer la administración de recursos (Seclen & Fabricio, 2023). Su aplicación en instituciones de educación básica y media ha mostrado beneficios en la identificación de patrones y tendencias que facilitan la implementación de estrategias para la mejora del rendimiento estudiantil (Villegas et al, 2020; Junior et al. 2024).

En Colombia, el Ministerio de Tecnologías de la Información y las Comunicaciones (MinTIC) promueve el uso de datos abiertos para generar conocimiento y mejorar los procesos en diversos sectores, incluyendo la educación (Ministerio de Tecnologías de la Información y las Comunicaciones, 2023). Sin embargo, aunque existen iniciativas gubernamentales que generan datos relevantes, su aprovechamiento aún es limitado. Un caso representativo es la estrategia “Evaluar para avanzar” (EPA), implementada por el Ministerio de Educación Nacional (MEN) desde 2021. Esta estrategia proporciona datos a nivel nacional, territorial e institucional sobre el desempeño de los estudiantes, con el propósito de apoyar la enseñanza y mejorar la calidad educativa. No obstante, los resultados de EPA aún no han sido articulados con otras fuentes de información para tomar decisiones estratégicas a nivel educativo.

A pesar de la disponibilidad de datos relacionados con EPA, su impacto en la mejora educativa sigue siendo limitado debido a la falta de integración con otras variables institucionales y a la ausencia de herramientas que permitan consolidar, relacionar y analizar esta información de manera efectiva. Esto impide una comprensión integral del proceso educativo, no solo en términos de resultados, sino también en su contexto (Villegas et al., 2020). Esta brecha dificulta que los directivos docentes y las secretarías de educación tomen decisiones informadas para fortalecer las políticas educativas y optimizar la gestión académica. Por ejemplo, esto impide la detección temprana de problemas en la calidad educativa o afecta la planificación de estrategias de mejora en instituciones de bajo desempeño.

El diseño de una arquitectura de inteligencia de negocios (BI) que integre los resultados de “Evaluar para avanzar” con las condiciones institucionales de los planteles educativos facilitará el análisis de relaciones entre variables y la identificación de patrones clave para fortalecer la toma de decisiones en políticas educativas (Villegas et al., 2020). En este sentido, esta monografía propone el diseño de una arquitectura BI aplicada a los centros educativos oficiales de Montería durante el periodo 2021-2022, con el objetivo de consolidar y analizar la información de manera efectiva, proporcionando una base sólida para la formulación de estrategias educativas informadas.

Para abordar este problema, se construye una arquitectura de BI tradicional de cuatro capas, que permita consolidar y analizar los datos de “Evaluar para avanzar” junto con información institucional de los centros educativos oficiales de Montería. La arquitectura empleará datos abiertos del Instituto Colombiano para la Evaluación de la Calidad de la Educación (ICFES) y datos anonimizados de la Secretaría de Educación de Montería. Se espera que la integración de estas fuentes de información brinde herramientas que faciliten la identificación de patrones, la generación de alertas tempranas

y la toma de decisiones basadas en evidencia, impactando de manera positiva la calidad educativa en el municipio y sirviendo como modelo replicable en otras regiones del país.

El presente trabajo de investigación se estructura de la siguiente manera: en la sección 1 se contextualiza el estudio; en la sección 2 se definen los objetivos generales y específicos; la sección 3 desarrolla la justificación del proyecto. Posteriormente, el marco teórico (sección 4) profundiza en los fundamentos conceptuales de la inteligencia de negocios (BI), su arquitectura (capas ETL, análisis y distribución), la estrategia “Evaluar para avanzar” (EPA) y su aplicación en el sector educativo, junto con métodos de analítica de datos. La sección 5 plantea la hipótesis, mientras que la sección 6 detalla las fuentes de datos y variables utilizadas. El diseño metodológico (sección 7) describe el enfoque, población, instrumentos, herramientas tecnológicas y técnicas de análisis (KPI y modelo de regresión). La sección 8 presenta el diseño de la arquitectura BI propuesta, desglosando cada capa (fuente de datos, ETL, análisis y distribución). Los resultados del trabajo de campo (sección 9) incluyen el procesamiento de datos, análisis de indicadores (desempeño, recursos humanos, equidad, socioeconómicos) y el modelo predictivo. Finalmente, las secciones 10 y 11 ofrecen directrices metodológicas, conclusiones, recomendaciones y líneas futuras de investigación.

2. Objetivos

2.1. Objetivo general

Proponer una arquitectura de inteligencia de negocios que permita integrar los resultados de la estrategia “Evaluar para avanzar” y las condiciones institucionales de los centros educativos oficiales de Montería durante el periodo 2021-2022 con el fin de revelar patrones que contribuyan a fortalecer la toma de decisiones en términos de políticas educativas.

2.2. Objetivos específicos

- Establecer referentes teóricos para diseñar una arquitectura de inteligencia de negocios que permita integrar los resultados de la estrategia “Evaluar para avanzar” y las condiciones institucionales de los centros educativos oficiales de Montería.
- Caracterizar fuentes de datos, y aplicar procesos de extracción, transformación y carga de datos de la estrategia “Evaluar para avanzar” y las condiciones de instituciones educativas oficiales de la ciudad de Montería.
- Configurar un sistema de almacenamiento de datos que consolide los resultados de la estrategia “Evaluar para avanzar” y las condiciones institucionales de los centros educativos oficiales de Montería.
- Generar visualizaciones de los resultados de la estrategia “Evaluar para avanzar” y las condiciones institucionales de los centros educativos oficiales de Montería.
- Proponer directrices metodológicas para diseñar arquitecturas de inteligencia de negocios para el sector educativo en Colombia.

3. Justificación

Dentro del marco de acción del Ministerio de Educación Nacional de Colombia, gestionado por el ICFES, se implementan diversas herramientas y estrategias diseñadas para respaldar y fortalecer los procesos educativos del país. Una de estas iniciativas es “Evaluar para avanzar” 3° a 11° (EPA), cuyo propósito principal es ofrecer herramientas de uso voluntario que apoyen a los docentes en sus procesos de enseñanza. Esta estrategia proporciona información valiosa para identificar niveles de desarrollo de competencias en las áreas evaluadas, diseñar estrategias de nivelación y fomentar el mejoramiento continuo de los procesos educativos (Icfes, 2023a). Adicionalmente, EPA permite monitorear el progreso de los estudiantes, lo que puede facilitar la implementación de acciones educativas tanto a nivel de aula como institucional. Sin embargo, esta herramienta, concebida como un insumo complementario, puede maximizar su impacto si se integra con información adicional, como las condiciones institucionales de los planteles educativos.

A pesar de su potencial, la implementación de EPA enfrenta desafíos, como la falta de análisis que relacionen sus resultados con factores institucionales y la carencia de herramientas tecnológicas para consolidar y contextualizar la información a nivel local. Estas limitaciones dificultan la toma de decisiones basadas en datos por parte de directivos docentes y secretarías de educación, quienes enfrentan retos en la priorización de recursos y la gestión de políticas educativas. Estudios recientes, como los del Laboratorio de Economía de la Educación de la Pontificia Universidad Javeriana (2022) y el ICFES (2023), destacan la necesidad de fortalecer la gestión educativa mediante herramientas tecnológicas que integren datos de diversas fuentes.

En este contexto, surge la propuesta de diseñar una arquitectura de inteligencia de negocios (BI) que consolide y analice datos de EPA y las condiciones institucionales de los planteles educativos oficiales de Montería. Esta arquitectura permitiría integrar y presentar

datos en visualizaciones dinámicas, accesibles para directivos docentes y secretarías de educación, actores clave en la toma de decisiones estratégicas y la gestión de recursos educativos. La propuesta proporcionaría indicadores clave y datos contextualizados para identificar brechas, monitorear el desempeño institucional y diseñar estrategias más efectivas. Aunque el enfoque principal son los directivos docentes y las secretarías de educación, otros actores también se beneficiarían. Los docentes podrían implementar mejoras en el aula, mientras que estudiantes y familias se verían impactados positivamente por una mejor calidad educativa. Los entes territoriales y diseñadores de políticas públicas obtendrían datos estructurados para formular y evaluar políticas, y los investigadores académicos podrían utilizar la arquitectura como modelo replicable y fuente de datos organizada.

El municipio de Montería se selecciona como caso de estudio debido a la disponibilidad de datos proporcionados por la Secretaría de Educación Municipal (SEM), lo que facilita el diseño, implementación y validación de la arquitectura. Esta elección permite evaluar cómo los datos de EPA, combinados con variables institucionales, pueden integrarse para generar información útil y procesable que respalde la toma de decisiones en niveles estratégicos. Aunque el modelo se enfoca en Montería, tiene potencial para ser adaptado y replicado en otras regiones del país, contribuyendo a una gestión educativa más efectiva y sostenible.

Desde una perspectiva académica, esta monografía se fundamenta en los conocimientos adquiridos durante la Maestría en Inteligencia de Negocios de la Universidad EAN y se alinea con las líneas de investigación en Ambientes de Aprendizaje, enfocadas en la Innovación y Pedagogía en Educación. El diseño de esta arquitectura de BI no solo responde a necesidades específicas, sino que también sienta las bases para el desarrollo de herramientas tecnológicas aplicadas al sector educativo colombiano.

4. Marco Teórico

4.1. Inteligencia de negocios.

La inteligencia de negocios (BI) se define como un conjunto de herramientas, técnicas y procesos esenciales para las organizaciones, permitiéndoles recopilar, analizar y visualizar datos con el propósito de tomar decisiones más informadas en cualquier ámbito empresarial (Salluco, 2021). Este concepto se fundamenta en la creciente importancia de los datos en la toma de decisiones, la necesidad de tomar decisiones adecuadas en un entorno competitivo. Además, es crucial destacar que la BI debe estar orientada a la acción, asegurando que la información derivada se traduzca en medidas concretas que beneficien a la organización, lo que la convierte en una herramienta esencial para mejorar el rendimiento empresarial. Esto implica la transformación de datos en información útil y conocimiento que respalda la elaboración de planes y estrategias. En este sentido, la BI permite integrar datos de diversas fuentes para facilitar el análisis y la toma de decisiones (Gómez-Duque et al., 2023; Mendoza-Rivera, 2022)

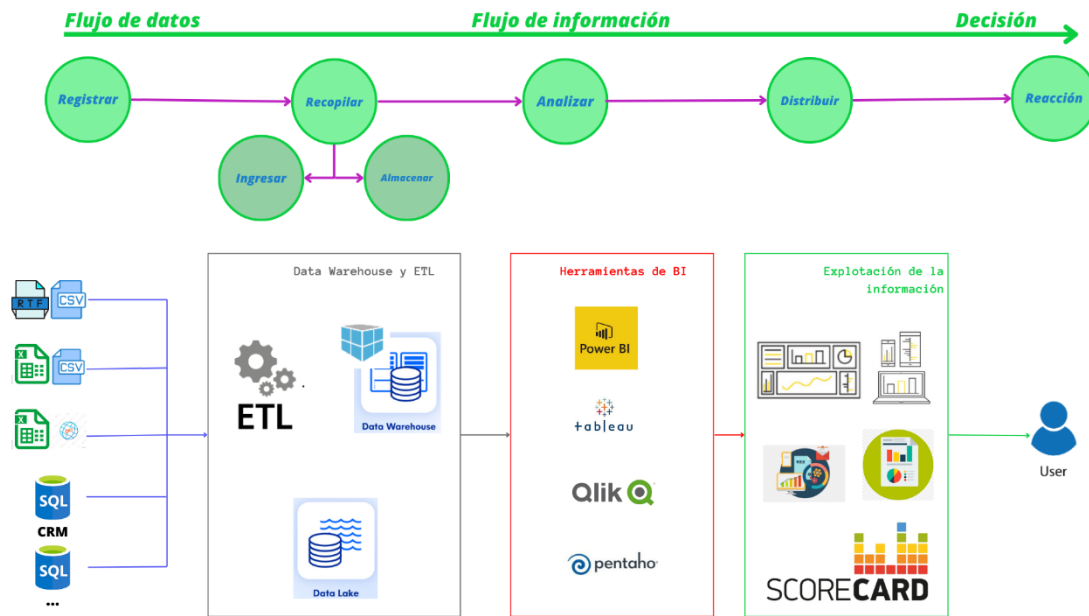
4.2. Arquitectura de inteligencia de negocios.

Ya se ha establecido como el BI combina arquitecturas, herramientas, bases de datos, herramientas analíticas y metodologías para permitir el acceso interactivo a los datos, su manipulación y el análisis para la toma de decisiones informadas (Sharda et al., 2018), donde el punto principal de convergencia entre los distintos autores es la arquitectura (Sorour et al., 2020; Velosa y Pabon, 2021; Caicedo, 2022), siendo esta un modelo de diseño básico para la aplicabilidad de una estrategia de BI. Sharda et al., (2018) indica que un sistema de BI consta de cuatro componentes principales: un almacén de datos (Data Warehouse, DW), herramientas de analítica empresarial, gestión de procesos

de negocio (Business Process Management, BPM) e interfaz de usuario. Se puede afirmar que la base de una arquitectura BI debe contener mínimo estos cuatro componentes o también denominados capas, de las cuales difieren según la necesidad de los requerimientos, abordando aspectos como: almacenamiento de datos, proceso de extracción, transformación y carga (ETL) y análisis de datos. Existen otros modelos de arquitectura de BI más específicos para aplicaciones o industrias determinadas, de las cuales destacamos el *modelo Kimball*, en el cual los datos se almacenan en un esquema dimensional (estrella), que es una estructura de datos simple y eficiente que facilita el análisis de datos. Los datos se dividen en dos tipos: hechos y dimensiones. Los hechos son las medidas que se van a analizar, como el número de ventas o el costo de los productos. Las dimensiones proporcionan contexto a los hechos, como el producto, el cliente o la fecha (Kimball & Ross, 2013, pág. 18). El proceso de extracción, transformación y carga (ETL) se utiliza para cargar los datos en el almacén de datos para el análisis de datos históricos (Inmon, 2013, pág. 16).

En línea con esta base, se reconoce que una arquitectura BI consta de capas y componentes interconectados para recopilar, almacenar, procesar y analizar datos (Caicedo, 2022). Este enfoque de múltiples capas, que asegura escalabilidad y calidad de datos, se adapta a las cambiantes demandas organizacionales. Rescatando y aplicando capas de la arquitectura BI de Salluco (2021), ajustada al sector público, se propone una estructura que abarca la recopilación de datos (ver Figura 1), almacenamiento eficiente, procesos ETL, herramientas de análisis empresarial y visualización.

Figura 1. Arquitectura de inteligencia de negocios



Fuente: Arquitectura de Inteligencia de Negocio en la Administración Pública, Salluco (2021).

Este esquema de arquitectura de BI aplicado por Salluco (2021), ofrece una estructura o arquitectura de cuatro (4) capas donde los datos se convierten en información y recursos de análisis. Cabe destacar que, para la aplicación de una arquitectura BI, se debe contextualizar la implementación (Romero, 2023), no siendo necesario especificarla como una capa dentro del modelo, pero siendo esencial para poder comprender la naturaleza del negocio, realizar un análisis de la situación actual efectiva y determinar los requerimientos de este (Herrera et al., 2022). Para el caso de esta monografía la naturaleza del negocio se centra en el sector educativo y la articulación de las pruebas EPA con las condiciones institucionales de los planteles educativos oficiales del municipio de Montería,

esta fase permite entender los datos que se van a utilizar, su disponibilidad y la forma de procesarlos, llevándonos a la formalización de las siguientes capas:

4.2.1. Capa de fuente de datos

Esta capa, según la perspectiva teórica (Salluco, 2021; Khatibi et al., 2020), se enfoca en la recopilación inicial de datos provenientes de diversas fuentes, es decir, todas aquellas fuentes de información como sistemas transaccionales y archivos. Para esta monografía, esta capa es de especial relevancia debido a la necesidad de recopilar información precisa y completa sobre las evaluaciones de la estrategia EPA y las condiciones institucionales. Aquí se pone énfasis en asegurar la obtención de datos mediante prácticas rigurosas que incluyen la verificación de integridad mediante procesos de validación y reconciliación.

4.2.2. Capa extracción, transformación y carga (ETL)

La capa de recopilación se posiciona como la fase inicial de extracción, donde los datos extraídos de diversas fuentes son sometidos a un proceso integral de limpieza y transformación para asegurar la calidad de la información. En este contexto, se emplean dos herramientas fundamentales de BI, la primera se refiere al proceso de Extracción, Transformación y Carga (ETL) la cual, según Medina (2023), constituye un conjunto de pasos para incorporar datos en un sistema de administración de bases de datos. Aquí, el enfoque inicia con la representación de los niveles de abstracción del proceso ETL. Se adopta un modelo lógico dimensional, con esquemas estrella, copo de nieve y una variante híbrida. Este último, un diseño híbrido, se esquematiza a partir de un modelo relacional normalizado, combinando elementos de ambos. El proceso ETL, siguiendo la descripción de López & Andrés (2023), implica la extracción de datos de una o varias fuentes, la

aplicación de procedimientos de transformación como operaciones aritméticas y lógicas, y finalmente, la carga de datos para establecer una fuente de verdad confiable y útil para procesos posteriores.

4.2.3. Capa de almacenamiento

En esta capa se encuentra el almacén de Datos (Data Warehouse - DWH), referido como el repositorio central para almacenar y organizar datos (Salluco, 2020), se aborda como un DWH no volátil y variable en el tiempo, según las características establecidas por Bill Inmon, (1996). La no volatilidad implica que, una vez almacenados, los datos no pueden ser modificados ni eliminados, convirtiéndose en información de solo lectura. Además, la capacidad de registrar y mantener un historial de cambios es esencial para la generación de informes históricos y el seguimiento de la evolución de los datos a lo largo del tiempo (Duplaa, 2022).

Esta capa, que abarca la recopilación, transformación y almacenamiento de datos, constituye un paso crucial para garantizar la consistencia y confiabilidad de la información que servirá como base para las fases subsiguientes del proyecto.

4.2.4. Capa de análisis

La Capa de Análisis marca la fase de exploración profunda de datos almacenados en el Data Warehouse (DWH). Aquí se aplican diversas técnicas y herramientas para extraer información valiosa con el objetivo de respaldar decisiones informadas. Durante el proceso de explotación de la información, se lleva a cabo una exploración exhaustiva utilizando las relaciones multidimensionales para acceder en tiempo real a información (utilizar de forma relacionada valores de distintas columnas y tablas). Estas relaciones son esenciales para examinar los datos desde diversas perspectivas, identificando relaciones

y patrones significativos. Además, algoritmos de minería de datos, empleando software especializado como Python, profundizan en la exploración para descubrir patrones y relaciones ocultas. Este enfoque integral garantiza una explotación máxima del potencial de los datos recopilados, estableciendo una sólida base para decisiones informadas (Duplaa, 2022; Khatibi et al., 2020; Sorour et al., 2020). Así mismo las herramientas de BI como Power BI son fundamentales en esta fase, destacando el uso de Power BI Desktop de Microsoft, una herramienta central para el análisis de datos. Power BI Desktop facilita la visualización en tableros de control interactivos, la creación de informes personalizados y la relación con lenguajes estadísticos, proporcionando funcionalidades esenciales para el análisis de datos en el contexto de la investigación (Duplaa, 2022).

Esta capa, centrada en la exploración y análisis profundo de datos, se conecta directamente con las capas anteriores, asegurando una transición fluida y una utilización efectiva de los datos recopilados en la toma de decisiones.

4.2.5. Capa de distribución

La fase de entrega de resultados y análisis a las partes interesadas es esencial para comunicar hallazgos de manera efectiva. Se generan informes y paneles interactivos que utilizan principalmente herramientas de visualización para crear gráficos, tableros de control y representaciones visuales de los datos, facilitando su comprensión. Los informes interactivos permiten una exploración detallada de los datos, vinculándose entre sí para actualizarse sincrónicamente al aplicar filtros. Según Khatibi et al. (2020), estos informes admiten funcionalidades como el desglose, que posibilita un análisis más detallado de datos, seleccionando una dimensión específica. Ofrecen diversas opciones de filtrado, como listas desplegadas, controles deslizantes y cuadros de búsqueda, permitiendo a los usuarios personalizar los resultados según sus necesidades.

En esta monografía, se emplea la herramienta de interacción y visualización de datos, Power BI, que proporciona diversas visualizaciones, desde datos numéricos hasta métricas y representaciones de texto (Microsoft, 2023). Se busca esquematizar tableros interactivos relacionando dimensiones asociadas para establecer informes ad hoc, manteniendo elementos definidos por Díaz (2010) para identificar KPIs clave por componente EPA, complementando la visualización con datos institucionales.

Esta capa cierra el ciclo al proporcionar a las partes interesadas una presentación clara y accesible de los resultados analíticos, asegurando que la información se utilice de manera efectiva en la toma de decisiones y la mejora continua.

4.3. Estrategia “Evaluar para avanzar” 3° a 11° (EPA)

La estrategia “Evaluar para avanzar” 3° a 11° es una iniciativa del gobierno nacional aplica a nivel nacional en todas las instituciones educativas que la adopten, a través del Instituto Colombiano para la Evaluación de la Educación (ICFES), se puede definir como un conjunto de herramientas de evaluación voluntarias destinadas a respaldar y acompañar el proceso de enseñanza de los docentes. Su objetivo principal es proporcionar información valiosa para diseñar estrategias educativas con un enfoque diagnóstico en las competencias y habilidades de los estudiantes desde los grados 3° a 11°, lo que a su vez ayuda a mejorar las estrategias de enseñanza y fortalecer la evaluación formativa (ICFES, 2023).

Dicha estrategia forma parte de la *Caja de Herramientas* del ICFES la cual se traduce en una serie de ayudas, gratuitas y de fácil acceso que le ayudarán a los y las estudiantes y docentes a entender, familiarizarse y aplicar metodologías de evaluación (Cuellar, 2022). Comenzó a funcionar en 2021 con una programación semestral para adaptarse a los dos calendarios educativos, conocidos como A y B, que operan en el país,

realizándose cuatro aplicaciones hasta la fecha con datos que se pueden interpretar (2021-1, 2021-2, 2022-1, 2022-2) (Laboratorio de Economía de la Educación (LEE) de la Pontificia Universidad Javeriana, 2022). La estrategia tiene dos formas de valoración, cualitativa y cuantitativa, utilizando guías estandarizadas de aplicación, indicando para la primera forma de valoración, una aplicación de preguntas abiertas a disposición de los docentes mediante cuadernillo de preguntas abiertas de competencias básicas y la segunda, valorada numéricamente mediante cuadernillos de opción múltiple con única respuesta los cuales se consolidan por estudiantes para su libre disposición que se publican semestre vencido, por ende se poseen cuatro aplicaciones de la estrategia de libre disposición en el repositorio de datos abiertos del ICFES *Data/CFES*, es decir (2021-1, 2021-2, 2022-1, 2022-2).

La estrategia tiene como propósito principal contribuir y monitorear las competencias de los estudiantes (ICFES, 2023, pág. 2). Al servir como herramienta de apoyo para el monitoreo y fortalecimiento del desarrollo de competencias, no solo aborda las áreas básicas valoradas, sino también aspectos relacionados con el desarrollo social y emocional, factores adicionales que influyen en el aprendizaje y la adaptación de los estudiantes ante situaciones de cambio, los instrumentos de implementación de la EPA están disponibles para aplicación en línea, sin conexión a internet y en físico. Estos cuadernillos se aplican para diferentes áreas (ver Figura 2) y grupos del área (relación de colores) de la cuales encontramos: Matemáticas (3º a 11º), Competencias Comunicativas en Lenguaje: Lectura (3º a 9º), Lectura Crítica (10º y 11º), Competencias Ciudadanas: Pensamiento Ciudadano y Ciencias Naturales y Educación Ambiental (5º a 9º), Sociales y Ciudadanas, y Ciencias Naturales (10º y 11º), cada uno con 20 preguntas, e Inglés (9º a 11º) que tiene 22 preguntas para grado 9º y 10º y 25 preguntas para grado 11º, de los cuales cada uno de estos instrumentos de valoración tiene una relación directa con los

estándares básicos de competencias visto a lo largo de cada periodo académico (ICFES, 2023, pág. 3).

Figura 2. Instrumentos de valoración de competencias básicas y cuestionarios auxiliares

Instrumentos de valoración	Grados								
	3°	4°	5°	6°	7°	8°	9°	10°	11°
Matemáticas	■	■	■	■	■	■	■	■	■
Competencias Comunicativas en Lenguaje: Lectura	■	■	■	■	■	■	■		
Lectura crítica								■	■
Ciencias Naturales y Educación Ambiental			■	■	■	■	■		
Ciencias Naturales								■	■
Competencias Ciudadanas: Pensamiento Ciudadano			■	■	■	■	■		
Sociales y ciudadanas								■	■
Inglés								■	■
Cuestionarios Auxiliares	■	■	■	■	■	■	■	■	■

Fuente: Instrumentos de valoración de competencias básicas y cuestionarios auxiliares (ICFES, 2023a).

Cabe destacar que para la valoración de las distintas áreas se remarca para todos los grados (3° a 11°) la aplicación de cuestionarios auxiliares, tomándose como una tercera forma de valoración los cuales, según el ICFES, pone a disposición de los docentes, especialmente a los directores de curso, un cuadernillo de cuestionarios auxiliares por ciclo educativo (básica primaria, básica secundaria y media), que busca identificar las creencias, actitudes y sentimientos de los estudiantes ante situaciones de cambio (ICFES, 2023, pág. 4), los cuales no tendrán trascendencia dentro de la monografía ya que sus resultados no están aún disponibles dentro del repositorio de datos abiertos (DataIcfes) y aunque su interpretación puede llevarse cuantitativamente (valores por agrupación de los estándares) su disposición es cualitativa mayormente.

Según lo establecido por el ICFES (2023), los agentes autorizados para acceder a la estrategia “Evaluar para avanzar” son todos los establecimientos educativos oficiales y

no oficiales del país. El proceso inicia con el Directivo docente del establecimiento educativo, quien debe ingresar a la plataforma electrónica <https://evaluarparaavanzar311.icfes.gov.co> (la cual solo se habilita en las ventanas de inscripción), realizar la validación de los estudiantes y preinscribir a los docentes de cada sede y jornada. Luego, los docentes acceden a la plataforma para finalizar su inscripción y asociar a sus estudiantes, permitiéndoles presentar los instrumentos de valoración en las modalidades que ofrece la estrategia (pág. 5).

4.4. Condiciones institucionales

Dentro del contexto educativo, especificando las instituciones educativas, el marco regulatorio colombiano, en cabeza del Ministerio de Educación Nacional (MEN) en base al Decreto 1330 de julio 25 de 2019 en su sección 3 ofrece una conceptualización de las condiciones institucionales las cuales se refieren a las características y elementos necesarios que existen dentro de la institución misma y que son fundamentales para facilitar y promover el desarrollo de las actividades formativas, académicas, docentes, científicas, culturales y de extensión. El propósito de considerar estas condiciones institucionales es fortalecer la institución en su totalidad y la comunidad académica: Todo proceso institucional debe llevarse a cabo en un marco de transparencia y buena gobernanza (Ministerio de Educación Nacional, 2019, Decreto 1330 de julio 25 de 2019).

En general, según dicha definición, las condiciones institucionales podrían clasificarse en dos categorías: condiciones estructurales y condiciones funcionales, siendo las condiciones estructurales aquellas que se refieren a la infraestructura, los recursos humanos y los recursos financieros de la institución educativa. Estas condiciones son fundamentales para el funcionamiento básico de la institución y para la prestación de los servicios educativos. Las condiciones funcionales son aquellas que se refieren a los

procesos, sistemas y mecanismos de gestión de la institución educativa. Estas condiciones son necesarias para garantizar la calidad de los servicios educativos como lo son: planeación, gestión, supervisión, evaluación, cultura organizacional, liderazgo y participación de la comunidad educativa.

Dichas condiciones revisten una importancia fundamental en el ámbito académico, tal como destaca Naranjo et al., (2021), donde señala que el rendimiento académico no se limita exclusivamente al desempeño individual del estudiantado; más bien, constituye un fenómeno complejo que abarca la institución educativa, su estructura, el entorno institucional, el cuerpo docente, los estudiantes, sus familias y la sociedad en su conjunto. Estos elementos interactúan de manera significativa y conjunta, y su influencia es determinante en el éxito o el fracaso estudiantil. Las valoraciones numéricas, consideradas como indicadores principales de rendimiento, reflejan un aspecto del rendimiento académico y están influenciadas por factores personales, de enseñanza, sociales e institucionales (Naranjo et al., 2021, pág. 4). En este contexto, es crucial reconocer que las condiciones institucionales, tanto estructurales como funcionales, desempeñan un papel esencial en la formación y el rendimiento de los estudiantes, convirtiéndose en un elemento central de interés para mejorar la calidad en el entorno educativo.

Para el caso de esta monografía se centra en la primera categoría la cual desde la perspectiva institucional resultan más fácil de medir y determinar, caracterizándola y relacionándola con los resultados institucionales.

4.4.1. Condiciones institucionales en relación con los resultados académicos

Los resultados académicos se suelen expresar a través de valoraciones numéricas que evalúan los distintos componentes asociados a las competencias correspondientes

(García, 2015, p. 4). Así mismo, este autor destaca la conclusión de diversos estudios que argumentan la influencia de factores diversos (institucionales, pedagógicos, psicológicos y sociodemográficos) en el rendimiento académico. Estos factores son esenciales para establecer un marco de acción en la mejora de la enseñanza. Naranjo et al, (2021) enfatizan en que el rendimiento académico va más allá del simple conocimiento, e incorpora factores personales, sociales e institucionales. Ejemplos claros de esta argumentación se encuentran mencionados por Rosero et al. (2021), donde se destaca la importancia de considerar diversos aspectos al analizar la complejidad del rendimiento académico, tal como el nivel educativo de los padres, el nivel de ingreso del hogar, el género, la naturaleza del colegio y su ubicación, las horas de dedicación al estudio y acceso a la tecnología, tienen mayor influencia sobre el rendimiento (Rosero et al., 2021, pp. 95). Así mismo se puede evidenciar en situaciones puntuales como cuando las familias con un nivel socioeconómico medio-alto suelen tener un cierto grado de holgura económica, lo que les permite estar alineadas con los proyectos educativos. Además, si enfrentan problemas o desafíos en el ámbito escolar, estas familias cuentan con recursos suficientes para influir en el contexto educativo (Cabrera, 2016, pp. 9). Así mismo encontramos como La experiencia de algunos estudiantes de nivel económico inoportuno constituye un principal indicador que afecta su capacidad para aprender y reflejar resultados óptimos (Valiente et al., 2014), dichos ejemplos reflejan contundentemente la injerencia de algunos factores dentro del rendimiento académico.

4.5. Inteligencia de negocios en un marco de aplicación al sector educativo y estrategias de evaluación y seguimiento.

4.5.1. Estudios en el ámbito internacional

Khatibi et al. (2020) señalan que la inteligencia empresarial (BI) desempeña un papel crucial a la hora de informar y dar forma a las estrategias educativas. Al integrar datos de diversas fuentes, los sistemas de BI pueden proporcionar a los tomadores de decisiones una visión integrada del sistema de educación superior, permitiéndoles monitorear, analizar y pronosticar indicadores clave. Esto puede ayudar a identificar tendencias, evaluar la eficacia de los programas educativos y tomar decisiones informadas para mejorar la calidad y eficiencia del sistema educativo.

Desde esa misma perspectiva, Sorour et al. (2020) exponen la importancia del BI para proporcionar sistemas analíticos que ayuden a los diferentes actores educativos a monitorear el aseguramiento de la calidad en las instituciones. Así mismo, estos autores enfatizan la necesidad de explorar el uso de datos de redes sociales en BI para capturar los sentimientos y perspectivas de los actores educativos en pro de mejorar su experiencia formativa. En esta vía, propone una arquitectura de BI para instituciones de educación superior, vinculando los principales elementos del sistema de garantía de calidad con el sistema de BI.

De igual forma se rescata lo indicado por Olivo (2023), donde revela que las soluciones de inteligencia de negocios se centran en abordar problemas relacionados con la gestión universitaria y el rendimiento académico. Este trabajo se enfoca en comprender cómo se utiliza la inteligencia de negocios en el contexto de la educación superior en América Latina, centrándose en la gestión y el rendimiento académico, utilizando una metodología que implica identificar las aplicaciones de la inteligencia de negocios en el sector educativo y conocer los países que están invirtiendo en investigación y desarrollo de inteligencia de negocios.

Chicaiza et al., (2022) proponen el desarrollo de una iniciativa tecnológica que incorpore herramientas de BI en el entorno académico de una institución de educación superior. La problemática abordada se centra en la escasa utilización de herramientas de BI en el ámbito educativo, subrayando la necesidad de integrar esta tecnología para obtener información en tiempo real sobre las operaciones académicas cruciales y facilitar la toma de decisiones fundamentadas. Según el autor, los estándares de calidad educativa establecidos por la Ley Orgánica de Educación Superior y sus respectivos indicadores respaldan a las instituciones educativas que buscan alcanzar la excelencia en la gestión de los procesos educativos.

Estos argumentos respaldan la aplicación BI en el sector educativo al destacar cómo el BI puede mejorar la toma de decisiones y la calidad educativa. Los autores anteriores enfatizan la importancia de integrar datos de diversas fuentes para monitorear y analizar indicadores clave, identificar tendencias, evaluar programas educativos y garantizar la calidad. Además, proponen arquitecturas de BI y metodologías que permiten a las instituciones de educación superior aprovechar el potencial del BI para mejorar la gestión y el rendimiento académico. En conjunto, estos argumentos respaldan la idea de que el BI puede ser una herramienta efectiva para impulsar la eficiencia y la calidad en el sector educativo al considerar variables y condiciones institucionales.

4.5.2. Estudios en el ámbito nacional

En el ámbito nacional, Vidales & Cleivert (2022) llevaron a cabo un estudio en Colombia aplicando herramientas de BI en el sector educativo. Utilizando la metodología de Ralph Kimball y la herramienta de visualización Microsoft Power BI, evaluaron el impacto de la implementación de BI en la Asociación Educativa Global Colombia S.A.S. Su investigación tuvo como objetivo mejorar la toma de decisiones en la gestión de

servicios, logrando predecir datos, evaluar la calidad del servicio y la productividad de manera satisfactoria.

De igual forma Valencia (2022), resalta la evolución del análisis de datos en la educación y señala que algunas instituciones, como la Institución Educativa Manuel María Mosquera en Puracé, no aprovechan completamente el potencial de los datos disponibles. Para abordar esto, implementaron una herramienta de inteligencia de negocios a través de un cuadro de mando, resultando en mejoras en los promedios académicos y los resultados del examen ICFES en los últimos dos años, dicho argumento apoya lo establecido por Gutiérrez et al., (2022), donde enfocan el análisis de los factores socioeconómicos relacionados con el rendimiento en las pruebas "Saber Pro" de 2016 para las Licenciaturas en Matemáticas en Colombia. Aunque no aplican directamente la metodología BI, su investigación destaca la importancia de evaluar estos factores y cómo la BI puede facilitar la recopilación, análisis y visualización de datos para tomar decisiones informadas y mejorar el rendimiento académico.

Por su parte, Vargas Agudelo et al., (2022), presentan un modelo de Inteligencia de Negocios para evaluar el impacto de la investigación y la innovación en programas de ingeniería en Colombia. Basado en experiencias de diversas Instituciones de Educación Superior y la metodología CRISP-DM, el modelo identifica variables e indicadores clave para comprender la contribución e impacto en el desarrollo académico y tecnológico.

Estos estudios respaldan la aplicación de BI en el contexto colombiano y el sector educativo, destacando su impacto positivo en la toma de decisiones y la mejora del rendimiento académico. Muestran cómo la implementación de BI ha permitido predecir datos, evaluar la calidad del servicio y la productividad en instituciones educativas. Además, resaltan la importancia de aprovechar al máximo el potencial de los datos y cómo el uso de BI ha contribuido a mejorar promedios académicos y resultados de exámenes

como el ICFES. La necesidad de comprender y evaluar factores socioeconómicos que influyen en el rendimiento académico respalda la aplicación de BI para recopilar y analizar estos datos.

4.5.3. Antecedentes de estudios analíticos de la EPA

En el ámbito de estrategias educativas centradas en la medición del rendimiento académico y la aplicación de metodologías de BI para analizar variables que pueden afectar directamente este rendimiento, se destaca la investigación realizada por el Laboratorio de Economía de la Educación (LEE) de la Pontificia Universidad Javeriana en relación con la estrategia APA. Este estudio, aunque no proporciona un marco específico de aplicación de BI utilizando datos de la estrategia, ofrece un antecedente claro que brinda una perspectiva valiosa sobre la importancia de las estrategias evaluativas en el ámbito educativo.

La investigación del LEE se alinea con la estrategia EPA en Colombia, la cual busca mejorar la calidad educativa mediante evaluaciones más flexibles y holísticas. Aunque la estrategia es novedosa y aún está en proceso de evaluación, se espera que los resultados de esta evaluación sean utilizados por diversas entidades, como docentes, colegios, Secretarías de Educación y el Gobierno Nacional. A pesar de algunas limitaciones, como la baja participación inicial de los estudiantes y ciertas inconsistencias en los datos, se anticipa que los resultados contribuirán al diseño de estrategias de nivelación académica y al monitoreo continuo del progreso de los estudiantes (Laboratorio de Economía de la Educación (LEE) de la Pontificia Universidad Javeriana, 2022).

Estos esfuerzos de evaluación y la posterior implementación de políticas efectivas son esenciales para mejorar la calidad de la educación en Colombia, ya que abordan directamente la pertinencia y el rendimiento académico en el contexto educativo actual. Es

importante destacar que, aunque el estudio es de naturaleza analítica, no incluye la implementación de soluciones específicas de BI.

La utilización de BI en este proceso tiene el potencial de enriquecer aún más esta mejora al permitir una evaluación más precisa y profunda de los datos relacionados con factores que inciden directamente en la pertinencia y el rendimiento académico. El análisis de datos a través de herramientas de BI puede proporcionar información valiosa para la toma de decisiones informadas y la formulación de políticas educativas que se adapten a las necesidades específicas de los estudiantes, identificando áreas de mejora y resaltando factores clave que impactan en la calidad de la educación.

4.6. Métodos de analítica de datos

Los métodos analíticos de datos son una forma de interpretar la información aplicando procesos de descripción, análisis e interpretación. Su desarrollo implica el uso de herramientas fundamentales como la estadística, algoritmos computacionales y sistemas de información (Cárdenas & León, 2023; Romero, 2023). En esta investigación, se adoptará el enfoque propuesto por Cárdenas & León (2023) y Sedkaoui (2018), quienes clasifican la analítica de datos en tres grupos: analítica descriptiva, analítica predictiva y analítica prescriptiva. Sin embargo, el enfoque principal se centrará en la analítica descriptiva y predictiva, con el objetivo de construir un cuadro de mando integral que integre indicadores clave y funciones de correlación para analizar la relación entre los resultados de la estrategia “Evaluar para avanzar” (EPA) y las condiciones institucionales.

Inicialmente, la analítica descriptiva constituye la fase inicial del análisis, donde se exploran datos históricos para identificar patrones de comportamiento en las variables de interés. En esta investigación, se realizará un análisis exploratorio de datos (EDA) que incluirá el diseño de indicadores claves descriptivos que integren los resultados de las

pruebas EPA con las condiciones institucionales. Estos indicadores permitirán visualizar y comparar el desempeño de los estudiantes en distintos contextos educativos, identificando factores institucionales que puedan influir en el rendimiento académico.

Posteriormente, la analítica predictiva se enfoca en el estudio de eventos pasados para crear modelos que anticipen tendencias o comportamientos futuros. En esta investigación, se utilizarán funciones de correlación para analizar la relación entre las variables independientes (condiciones institucionales) y la variable dependiente (resultados de la EPA), esto con el fin de calcular numéricamente el grado del rendimiento académico de los estudiantes en función de las diferentes condiciones institucionales.

Finalmente, el cuadro de mando será la herramienta central para integrar y visualizar los resultados del análisis, donde se incluyen: visualización de indicadores clave, funciones de correlación y análisis exploratorio

5. Hipótesis

El diseño de una arquitectura de BI para integrar los resultados de la estrategia “Evaluar para avanzar” y las condiciones institucionales es una alternativa para identificar patrones, examinar relaciones entre variables y generar visualizaciones que permitan fortalecer la toma de decisiones en términos de políticas educativas en el municipio de Montería.

6. Fuente de datos y Variables

6.1. Fuente de datos

Para el desarrollo de esta monografía, se han utilizado diversas fuentes de datos provenientes de entidades oficiales, con el propósito de consolidar una arquitectura de inteligencia de negocios aplicada al sector educativo en Montería. Las principales fuentes de información incluyen el reporte de matrículas anonimizadas de los años 2021-2022, datos de las instituciones educativas oficiales de Montería con información sobre la planta docente y administrativa de la Secretaría de Educación del municipio, registros geográficos de municipios y departamentos del DANE, y finalmente, los resultados de la estrategia “Evaluar para avanzar” (EPA) provistos por DataIcfes.

El reporte de matrículas anonimizadas de instituciones educativas oficiales de Montería 2021 – 2022, proporcionado por la Secretaría de Educación de Montería en formato Excel, incluye datos identificados con un código único asignado por la entidad, permitiendo así la trazabilidad de la información sin comprometer la privacidad de los estudiantes. Entre las variables registradas se encuentran la edad, género, grado, grupo, calendario académico, institución educativa, jornada y modelo educativo, estrato socioeconómico, apoyo académico recibido, condición de discapacidad y asistencia a programas de ayuda pedagógica. Esta fuente de datos permite caracterizar la población estudiantil y analizar su distribución dentro del sistema educativo del municipio.

Por otra parte, la base de datos de las instituciones educativas oficiales proporciona información estructurada sobre las instituciones educativas de Montería, identificándolas mediante su nombre y código DANE. Adicionalmente, se incluyen datos sobre sus distintas sedes, lo que facilita la integración y análisis de información académica y administrativa dentro de la arquitectura de inteligencia de negocios.

En cuanto a la planta docente y administrativa, la Secretaría de Educación de Montería dispone de datos institucionales correspondientes al personal vinculado al sector educativo durante 2012- 2022. Esta fuente de información detalla variables como la edad de los docentes, la institución educativa a la que pertenecen, el número de estudiantes asignados por docente, el escalafón docente, el tipo de contrato, el cargo ocupado, el área académica, el nivel educativo impartido, el género del empleado, el tiempo de servicio, la ciudad de residencia y la situación laboral. Estos datos permiten analizar la distribución del recurso humano en el sector educativo, identificando posibles correlaciones entre la estructura de la planta docente y los resultados educativos. Adicionalmente, se cuenta con datos geográficos de municipios y departamentos, una fuente transversal que facilita la vinculación geográfica de la información educativa. Este conjunto de datos, disponible en los registros abiertos del DANE, contiene el listado de municipios y departamentos de Colombia junto con sus respectivos códigos DANE. La información es extraída en formato XLSX desde la base de datos del gobierno nacional y posteriormente almacenada y procesada en archivos separados por comas (CSV). Su inclusión dentro del análisis permite relacionar la información educativa con la división territorial del país, lo que posibilita estudios comparativos y geoespaciales en el sector educativo. Finalmente, los datos de la estrategia “Evaluar para avanzar” (EPA) corresponden a los reportes de resultados de las evaluaciones aplicadas a nivel nacional durante 2021-2022. Esta fuente de información, disponible a través de Data Icfes (<https://www.icfes.gov.co/investigaciones/data-icfes/>) en formato CSV, recoge datos institucionales como el nombre de la institución educativa, su código DANE y los grados evaluados. Además, registra los resultados de la valoración en cada aplicación de la estrategia, desglosados por componentes, competencias y áreas de conocimiento, incluyendo el número y porcentaje de aciertos y desaciertos. La inclusión de estos datos

dentro del sistema de inteligencia de negocios permite analizar el desempeño académico de los estudiantes y relacionarlos con las condiciones institucionales de los centros educativos.

6.2. Variables

A continuación, se detallan las variables que hacen parte de este trabajo. En la Tabla 1 y Tabla 2 se describen las variables teniendo en cuenta la fuente de donde provienen, el nombre de la variable, la descripción, la unidad de medida, la naturaleza, la escala y el formato del archivo donde se ubica.

Tabla 1. Variables de la estrategia EPA

SIGLA	VARIABLE	DESCRIPCIÓN	TIPO DE VARIABLE	MEDIDA
RC	PROMEDIO DESEMPEÑO RESPUESTAS CORRECTAS POR INSTITUCION	Promedio del número de respuestas correctas por estudiante en la prueba EPA, calculado a nivel de institución educativa. Refleja el desempeño general de los estudiantes en la prueba.	Cuantitativa Continua	1-100

Fuente: Elaboración propia

Tabla 2. Variables de las condiciones institucionales

SIGLA	VARIABLE	DESCRIPCIÓN	TIPO DE VARIABLE	MEDIDA
IE	CODIGO INSTITUCIONES EDUCATIVAS	Código único asignado por el DANE para identificar cada institución educativa.	Cualitativa nominal	
NE	CONTEO DE ESTUDIANTES POR INSTITUCION	Número total de estudiantes matriculados en una institución educativa.	Cuantitativa Discreta	0-2000
ZOU	CONTEO DE ESTUDIANTES POR ZONA URBANA	Número de estudiantes ubicados en zonas urbanas por institución.	Cuantitativa Discreta	0-2000
ZOR	CONTEO DE ESTUDIANTES POR ZONA RURAL	Número de estudiantes ubicados en zonas rurales por institución.	Cuantitativa Discreta	0-2000
EM	CONTEO DE ESTUDIANTES MASCULINOS	Número de estudiantes de género masculino matriculados por institución.	Cuantitativa Discreta	0-2000

ARQUITECTURA DE INTELIGENCIA DE NEGOCIOS PARA INTEGRAR RESULTADOS DE LA ESTRATEGIA “EVALUAR PARA AVANZAR” Y LAS CONDICIONES INSTITUCIONALES DE PLANTELES EDUCATIVOS OFICIALES DEL MUNICIPIO DE MONTERÍA

EF	CONTEO DE ESTUDIANTES FEMENINOS	Número de estudiantes de género femenino matriculados por institución.	Cuantitativa Discreta	0-2000
DI	CONTEO ESTUDIANTES CON DISCAPACIDAD	Número de estudiantes con discapacidad matriculados por institución.	Cuantitativa Discreta	0-500
AA	CONTEO ESTUDIANTES CON APOYO ACADÉMICO ESPECIAL	Número de estudiantes que reciben apoyo académico especial por institución.	Cuantitativa Discreta	0-100
ES13	CONTEO ESTUDIANTES DE ESTRATO 1-3	Número de estudiantes que se encuentran entre los estratos 1 al 3 por institución.	Cuantitativa Discreta	0-2000
ES46	CONTEO ESTUDIANTES DE ESTRATO 4-6	Número de estudiantes que se encuentran entre los estratos 4 al 6 por institución.	Cuantitativa Discreta	0-2000
DS	CONTEO DOCENTES SECUNDARIA	Número de docentes que enseñan en el nivel de secundaria por institución.	Cuantitativa Discreta	0-50
DD	CONTEO DOCENTES DIRECTIVOS	Número de directivos docentes (rectores, coordinadores, etc.) por institución.	Cuantitativa Discreta	0-50
TS	TIEMPO SERVICIO	Suma total del tiempo de servicio (en años) de los docentes por institución.	Cuantitativa Continua	0-30
ED30	CONTEO DOCENTES EN EDAD 18-30	Número de docentes con edades entre 18 y 30 años por institución.	Cuantitativa Discreta	0-50
ED40	CONTEO DOCENTES EN EDAD 31-40	Número de docentes con edades entre 31 y 40 años por institución.	Cuantitativa Discreta	0-50
ED50	CONTEO DOCENTES EN EDAD 41-50	Número de docentes con edades entre 41 y 50 años por institución.	Cuantitativa Discreta	0-50
ED+	CONTEO DOCENTES EN EDAD 51+	Número de docentes con edades de 51 años o más por institución.	Cuantitativa Discreta	0-50
DPR	CONTEO DOCENTES PROPIEDAD	Número de docentes en propiedad (nombrados en propiedad) por institución.	Cuantitativa Discreta	0-50
DPD	CONTEO DOCENTES PROVISIÓN DEFINITIVA	Número de docentes en provisión definitiva por institución.	Cuantitativa Discreta	0-50
DPT	CONTEO DOCENTES PROVISIÓN TEMPORAL	Número de docentes en provisión temporal por institución.	Cuantitativa Discreta	0-50

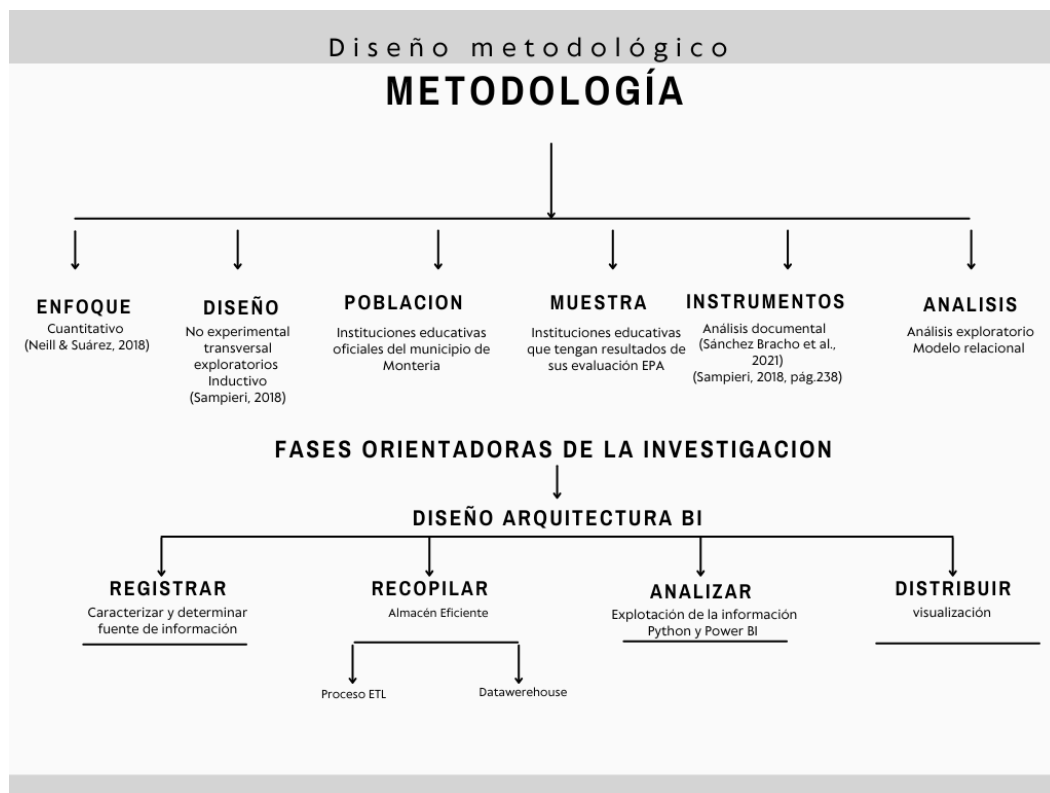
Fuente: Elaboración propia

7. Diseño metodológico

Esta investigación se enfoca en una investigación de tipo aplicada, en donde se tomarán datos disponibles para realizar un modelo que integre resultados de estrategias educativas (EPA) del sector oficial, con las condiciones institucionales de cada entidad (IE), para estructurar la aplicación de un diseño de arquitectura BI a un caso educativo del sector público del municipio de Montería.

Es así como la Figura 3, muestra la estructura del diseño metodológico aplicado a esta investigación y las fases orientadoras de la misma.

Figura 3. Diseño metodológico.



Fuente: Elaboración propia

7.1. Enfoque de la investigación

La presente investigación adopta un enfoque cuantitativo y descriptivo, dado que busca analizar, integrar y visualizar datos provenientes de la estrategia “Evaluar para avanzar” (EPA) junto con información institucional de los centros educativos oficiales de Montería. La utilización de un enfoque cuantitativo permite estructurar, procesar y analizar grandes volúmenes de datos de manera objetiva, facilitando la identificación de patrones, tendencias y correlaciones en el desempeño académico y las condiciones institucionales.

Además, el estudio tiene un carácter aplicado, pues su finalidad es diseñar una arquitectura de inteligencia de negocios (BI) que optimice el análisis y la gestión de datos en el sector educativo. Al estructurar y visualizar información relevante, se busca generar insumos útiles para la toma de decisiones estratégicas en el ámbito educativo.

7.2. Alcance de la Investigación de la investigación

El estudio se enmarca en un diseño no experimental, transversal y descriptivo, dado que se analizarán datos históricos correspondientes al periodo 2021-2022 sin manipulación de variables. Su alcance es descriptivo y propositivo, ya que busca caracterizar la información disponible en la estrategia EPA y en las bases de datos institucionales, para posteriormente proponer un modelo de integración basado en BI.

En términos geográficos, la investigación se centra en los centros educativos oficiales de Montería, considerando la información proporcionada por la Secretaría de Educación Municipal y el Instituto Colombiano para la Evaluación de la Calidad de la Educación (ICFES). Sin embargo, el modelo propuesto tiene potencial de replicabilidad en otros contextos educativos del país.

7.3. Población y muestra

La población del estudio está compuesta por las instituciones educativas oficiales de Montería, que han participado en la estrategia “Evaluar para avanzar” durante los años 2021 y 2022. La selección de la muestra se realiza bajo un criterio no probabilístico y por conveniencia, considerando los centros educativos que cuenten con datos disponibles tanto de EPA como de otras fuentes institucionales, garantizando la integridad de la información para el análisis.

La muestra específica incluirá aquellos centros educativos cuyos datos puedan ser consolidados en la arquitectura de BI propuesta, priorizando la disponibilidad y calidad de los registros en la Secretaría de Educación de Montería y el ICFES.

7.4. Instrumentos

La recolección de datos en investigación puede ofrecer una gran variedad de técnicas e instrumentos, independientemente de si el enfoque es cualitativo o cuantitativo. Lo crucial en una investigación reside en delimitar qué aspectos son medibles y cómo obtener esos datos (Sampieri, 2018). Este estudio, se centró en la recopilación y organización de datos proporcionados por fuentes oficiales. Se consultaron fuentes oficiales para obtener datos sobre los resultados de la estrategia EPA disponibles en la base de datos pública del ICFES conocida como DATAICFES, así mismo para la consulta de las variables institucionales relacionadas con el consolidado de matrícula por estudiante y la planta docente municipal de Montería, se realizó una solicitud por escrito a la Secretaría de Educación de Montería para acceder a los informes y registros oficiales en formatos estructurados, como archivos CSV y Excel (xlsx), los cuales fueron proporcionados por la entidad.

7.5. Herramientas tecnológicas para la construcción de la arquitectura de BI

La construcción de una arquitectura de BI requiere la selección de herramientas tecnológicas que permitan la extracción, transformación y carga (ETL) de datos, su almacenamiento estructurado, así como su visualización y análisis. En esta sección, se presenta una revisión de las tecnologías disponibles, sus ventajas y desventajas.

Para la construcción de la arquitectura de BI, se evaluaron diversas herramientas tecnológicas, considerando su capacidad para manejar grandes volúmenes de datos, su flexibilidad, su facilidad de uso y su integración con otras plataformas. La Tabla 3 presenta un cuadro comparativo de las herramientas más relevantes.

Tabla 3. Tecnologías para la construcción de una arquitectura BI.

Herramienta	Ventajas	Desventajas	Uso
Python	Amplia gama de bibliotecas para ETL (pandas, NumPy, SQLAlchemy).	-Curva de aprendizaje para usuarios no técnicos.	ETL
	Flexibilidad para manipular y analizar datos.	Requiere conocimientos de programación.	
	Integración con múltiples fuentes de datos.		
SQL Server	Potente para el manejo de bases de datos relacionales.	Menor flexibilidad para análisis avanzados y visualizaciones.	Almacenamiento
	Escalable y robusto para grandes volúmenes de datos.	Requiere conocimientos avanzados de SQL.	

Power BI	Interfaz intuitiva y fácil de usar.	Limitaciones en el manejo de grandes volúmenes de datos sin optimización.	Visualización
	Amplia gama de visualizaciones interactivas.	Costos asociados para versiones premium.	
	Integración con múltiples fuentes de datos y herramientas de BI.		
Tableau	Visualizaciones avanzadas y altamente personalizables.	Costo elevado en comparación con otras herramientas.	Visualización
	Fácil integración con bases de datos y herramientas de análisis.	Menor flexibilidad para transformaciones complejas de datos.	
Apache Spark	Ideal para el procesamiento de grandes volúmenes de datos en tiempo real.	Configuración y mantenimiento complejos.	ETL
	Escalabilidad y alto rendimiento.	Curva de aprendizaje pronunciada.	
	Automatización de flujos de datos sin programación.		
Talend	Plataforma robusta con gran cantidad de conectores.	Licencia costosa en versiones empresariales.	ETL
	Soporte para ETL en la nube y local.	Puede ser innecesariamente complejo para pequeños proyectos.	
Amazon Redshift	Escalable y optimizado para grandes volúmenes de datos.	Costos asociados al uso en la nube.	Almacenamiento

Google BigQuery	- Procesamiento rápido de grandes volúmenes de datos.	Costos asociados al uso en la nube.	Almacenamiento
QlikView	Integración con herramientas de BI y análisis.	Dependencia de la infraestructura de Google Cloud.	Visualización
	Visualizaciones interactivas y personalizables.	Costo elevado en comparación con otras herramientas.	
	Fácil integración con múltiples fuentes de datos.	Menor flexibilidad para transformaciones complejas de datos.	

Fuente: Elaboración propia

En la arquitectura propuesta, Python se utilizará para la extracción, transformación y carga de datos desde diversas fuentes hacia SQL Server, que actuará como la base de datos centralizada para el almacenamiento. Posteriormente, Power BI se conectará a SQL Server para generar tablero de controls y reportes interactivos que permitan el análisis de la información de manera eficiente. Esta combinación garantiza una solución escalable, flexible y de alto impacto para la toma de decisiones basada en datos.

7.6. Técnicas para el análisis de la información

Dentro de las técnicas de análisis de la información, se ha determinado dos aspectos fundamentales, el primero es el cumplimiento del enfoque de la investigación y sus diferentes objetivos mediante análisis exploratorio de datos mediante la implementación de indicadores claves (KPI), acompañado de procedimientos descrito por Hernández Sampieri, (2006) donde se encuentran los modelos de regresión lineal para examinar relaciones entre variables.

7.6.1. Indicadores Claves de Desempeño (KPI)

La creación de indicadores clave es un proceso que depende en gran medida del contexto en el que se desarrolla la investigación y de las necesidades específicas del estudio. Como señala Valladares y Posso (2022), los indicadores no siguen un único modelo, sino que se diseñan de distintas maneras según las características del entorno y los objetivos de la investigación. En este sentido, los indicadores no son universales, sino que deben diseñarse para reflejar las dimensiones y variables relevantes del fenómeno estudiado. En el marco de esta investigación, los indicadores clave se diseñarán para integrar los resultados de la estrategia EPA con las condiciones institucionales, con el fin de proporcionar una visión holística del sistema educativo. Estos indicadores se clasificarán en varios tipos, dependiendo de su enfoque y utilidad tal como los ilustra la Tabla 4, Tabla 5, Tabla 6 y Tabla 7.

Los indicadores que se presentan a continuación son de autoría propia y están inspirados en los criterios metodológicos de otros estudios educativos que mencionan la importancia de crear indicadores para monitorear y relacionar la actividad educativa. Por ejemplo, Agyei et al. (2018) destacan la relevancia de los indicadores para medir la desigualdad en la infraestructura educativa y su impacto en el rendimiento académico, demostrando que las condiciones institucionales, como la distribución de recursos, condicionan los resultados educativos a nivel distrital. Asimismo, Hanushek y Woessmann (2021) enfatizan que las habilidades medidas a través de pruebas estandarizadas, debe monitorizarse porque son un componente clave del capital de conocimiento de una nación y tienen un impacto significativo en el crecimiento económico a largo plazo.

En el ámbito del desempeño educativo, Rodríguez Vargas et al. (2023) proponen una metodología para diseñar indicadores de desempeño en contextos técnicos y profesionales, resaltando la necesidad de medir dimensiones y subdimensiones

específicas alineadas con las exigencias del modelo profesional. Por otro lado, Marciniak y Cáliz Rivera (2021) desarrollan un sistema de indicadores para evaluar la calidad de los materiales didácticos en entornos de educación en línea, lo que refuerza la importancia de diseñar indicadores validados que guíen procesos educativos clave.

Finalmente, en el ámbito socioeconómico, Liu et al. (2022) demuestran que el estatus socioeconómico (SES) está significativamente correlacionado con el rendimiento académico, lo que subraya la necesidad de monitorear indicadores socioeconómicos para abordar desigualdades en los sistemas educativos. Estos estudios, entre otros, respaldan el diseño de indicadores que no solo midan resultados académicos, sino que también integren condiciones institucionales y socioeconómicas para ofrecer una evaluación más completa y equitativa del sistema educativo.

Tabla 4. Indicadores de desempeño académico

Nombre del Indicador	Fórmula	Descripción	Escala	Interpretación
Desempeño académico promedio por institución educativa	Suma total de respuestas correctas por institución / Número total de estudiantes que respondieron en la institución	Mide el rendimiento académico promedio de cada institución en las pruebas EPA.	0 a 1 o 0% a 100%	Un valor más alto indica mejor rendimiento institucional. Ideal para identificar instituciones con mejor o peor desempeño.
Desempeño académico por instrumento y competencia	Suma total de respuestas correctas por competencia / Número de estudiantes que respondieron esa competencia	Evalúa el rendimiento promedio por área evaluada (matemáticas, inglés, etc.) y sus competencias específicas.	0 a 1 o 0% a 100%	Permite comparar el rendimiento por áreas y competencias específicas, identificando fortalezas y debilidades temáticas.

Desempeño académico promedio por grado escolar	Suma de respuestas correctas por grado / Número total de estudiantes del grado	Mide el promedio de rendimiento académico según el nivel educativo (3° a 11°).	0 a 1 o 0% a 100%	Permite observar el progreso o retroceso del rendimiento a través de los diferentes grados escolares.
Índice de desempeño por jornada escolar	Suma de respuestas correctas por jornada / Número de estudiantes por jornada	Mide el promedio de desempeño según la jornada (mañana, tarde, única).	0 a 1 o 0% a 100%	Ayuda a identificar si existen diferencias significativas en el rendimiento según la jornada escolar.
Desempeño promedio por zona geográfica	Suma total de respuestas correctas por zona / Número total de estudiantes en la zona	Compara el rendimiento académico entre zonas urbanas y rurales.	0 a 1 o 0% a 100%	Muestra brechas entre zonas geográficas, útiles para decisiones de política educativa.
Desempeño promedio por género	Suma total de respuestas correctas por género / Número de estudiantes por género	Compara el promedio de desempeño entre hombres y mujeres.	0 a 1 o 0% a 100%	Permite identificar desigualdades de género en los resultados académicos.
Distribución de estudiantes por rango de rendimiento	Clasificación de los estudiantes según: Muy bajo (<25%), Bajo (25–50%), Medio (50–75%), Alto (>75%)	Agrupar a los estudiantes según su nivel de desempeño.	% de estudiantes por rango	Ayuda a visualizar la proporción de estudiantes en cada nivel de rendimiento y focalizar intervenciones.
Índice de brecha de género por grado escolar	Promedio de respuestas correctas hombres – promedio de respuestas correctas mujeres (por grado)	Mide la diferencia en el rendimiento académico entre hombres y mujeres en cada grado.	Valores positivos, negativos o 0	Positivo: mejor desempeño masculino. Negativo: mejor desempeño femenino. Cero: equidad de género.

Fuente: Los indicadores fueron contruidos por el autor con base en criterios metodológicos tomados de referentes como Agyei et al. (2018), Marciniak y Cáliz Rivera (2021), Liu et al. (2022), Araiza Lozano, M. Á. (2021), Menendez, L. (2022), García-Herrero et al., (2024) y Ortigoza et al., (2024).

Tabla 5. Indicadores de recursos humanos.

Nombre del Indicador	Fórmula	Descripción	Escala	Interpretación
Relación docente/estudiante por área y grado	Promedio de estudiantes por institución / Promedio de docentes por institución (en el área y grado correspondientes)	Mide la cantidad promedio de estudiantes asignados a un docente, desglosado por área y grado escolar.	Número entero o decimal	Un valor más alto indica mayor carga docente. Ayuda a identificar necesidades de contratación o redistribución del personal.
% de docentes por tipo de contrato	(Número de docentes por tipo de contrato / Total de docentes) × 100	Muestra la proporción de docentes con contrato en propiedad, provisionales u otros.	Porcentaje (%)	Permite evaluar la estabilidad laboral del cuerpo docente y la dependencia de formas contractuales temporales.
% de distribución de los docentes por rango de edad	(Número de docentes en cada rango de edad / Total de docentes por institución) × 100	Indica la proporción de docentes según grupos etarios: 18–30, 31–40, 41–50, 51 o más.	Porcentaje (%)	Ayuda a identificar el grado de envejecimiento o renovación generacional del equipo docente. Útil para prever relevos y jubilaciones.
Índice de impacto de la jornada académica en el desempeño	Promedio de respuestas correctas por jornada / Total de estudiantes por jornada	Mide el rendimiento académico promedio por tipo de jornada escolar (mañana, tarde, única).	Valor normalizado (0 a 1 o % promedio)	Un valor alto sugiere que cierta jornada favorece el aprendizaje. Permite optimizar la organización horaria institucional.
Desempeño académico promedio docente por instrumento	Suma total de rendimiento promedio de los estudiantes por docente / Número de docentes (por instrumento)	Calcula el promedio de desempeño docente en función del rendimiento de sus estudiantes por área evaluada.	0 a 1 o 0% a 100%	Identifica fortalezas o debilidades en el desempeño asociado a cada área de evaluación (matemáticas, inglés, lectura, etc.).

Distribución de docentes por área académica	(Número de docentes por área / Total de docentes) × 100	Muestra la proporción de docentes distribuidos en diferentes áreas académicas (lenguaje, ciencias, matemáticas, etc.).	Porcentaje (%)	Permite analizar la conformación del equipo docente según las áreas del conocimiento. Es útil para análisis de cobertura y especialización.
--	---	--	----------------	---

Fuente: Los indicadores fueron construidos por el autor con base en criterios metodológicos tomados de referentes como Agyei et al. (2018), Marciniak y Cáliz Rivera (2021), Liu et al. (2022), Araiza Lozano, M. Á. (2021), Menendez, L. (2022), García-Herrero et al., (2024) y Ortigoza et al., (2024).

Tabla 6. Indicadores de equidad educativa.

Nombre del Indicador	Fórmula	Descripción	Escala	Interpretación
Índice de diversidad estudiantil por institución	(Número de estudiantes con condición especial en la institución / Total de estudiantes en la institución) × 100	Mide el grado de inclusión en cada institución educativa, reflejando la proporción de estudiantes con condiciones especiales (SRPA, discapacidad física o cognitiva, apoyo especial).	Porcentaje (%)	Un valor más alto indica mayor diversidad estudiantil y la posible necesidad de recursos especializados, infraestructura inclusiva y estrategias pedagógicas diferenciadas.
Proporción de estudiantes por tipo de condición especial	Número de estudiantes por tipo de condición / Total de estudiantes con condición especial	Determina la distribución porcentual de los distintos tipos de condiciones especiales (ninguna, SRPA, discapacidad física, discapacidad cognitiva, apoyo especial) entre los estudiantes identificados con diversidad.	Porcentaje (%)	Permite analizar qué tipo de condición especial es más frecuente en la población estudiantil y orientar la asignación de apoyos específicos según las necesidades predominantes.
Desempeño académico promedio por tipo de condición especial	Sumatoria de respuestas	Compara el promedio de desempeño académico entre	De 0 a 1 o de 0% a 100%	Permite identificar brechas de rendimiento

correctas por grupo de condición / Total de estudiantes por grupo	estudiantes con y sin condiciones especiales. Puede desglosarse por instrumento, componente o competencia.	asociadas a condiciones especiales, lo cual es clave para implementar estrategias de inclusión educativa efectivas.
---	--	---

Fuente: Los indicadores fueron contruidos por el autor con base en criterios metodológicos tomados de referentes como Agyei et al. (2018), Marciniak y Cáliz Rivera (2021), Liu et al. (2022), Araiza Lozano, M. Á. (2021), Menendez, L. (2022), García-Herrero et al., (2024) y Ortigoza et al., (2024).

Tabla 7. Indicadores socioeconómicos.

Nombre del Indicador	Fórmula	Descripción	Escala	Interpretación
Distribución de estudiantes por estrato socioeconómico	$(\text{Número de estudiantes por estrato} / \text{Total de estudiantes}) \times 100$	Mide la proporción de estudiantes en cada nivel socioeconómico (estratos 1, 2, 3 y 4) dentro de una institución o conjunto de instituciones.	Porcentaje (%)	Permite caracterizar la composición socioeconómica de la población estudiantil. Un mayor porcentaje en estratos bajos puede evidenciar contextos de vulnerabilidad.
Desempeño académico promedio por estrato socioeconómico	$\text{Suma de respuestas correctas por estrato} / \text{Número de estudiantes en el estrato}$	Calcula el rendimiento académico promedio de los estudiantes agrupados según su nivel socioeconómico.	Escala de 0 a 1 o 0% a 100%	Permite identificar brechas de desempeño asociadas al nivel socioeconómico, lo cual es clave para implementar estrategias de equidad educativa.
Índice de concentración institucional por estrato	$\text{Número de estudiantes del estrato predominante en la institución} / \text{Total de estudiantes en la institución}$	Mide qué tan concentrada está una institución en un solo nivel socioeconómico.	Porcentaje (%)	Un valor alto indica que la mayoría de los estudiantes pertenecen al mismo estrato, lo que puede reflejar segregación social o desigualdad territorial.

Fuente: Los indicadores fueron contruidos por el autor con base en criterios metodológicos tomados de referentes como Agyei et al. (2018), Marciniak y Cáliz Rivera (2021), Liu et al. (2022), Araiza Lozano, M. Á. (2021), Menendez, L. (2022), García-Herrero et al., (2024) y Ortigoza et al., (2024).

Referente a los indicadores de relación, es necesario explicar el cálculo de estos indicadores difiere de los anteriores, debido a la complejidad de su construcción. Ya que estos combinan variables contextuales con resultados académicos mediante un modelo de clasificación y normalización, que permite establecer relaciones comparativas entre instituciones educativas.

Para calcular el indicador, primero se determinan las medianas del desempeño académico y de la condición institucional analizada (como número de docentes, aulas, etc.), lo cual permite establecer un punto de referencia ya que representa el valor central de cada variable, lo que permite clasificar a las instituciones en relación con su posición respecto al conjunto. A diferencia del promedio, la mediana no se ve afectada por valores extremos o atípicos, lo que la convierte en una medida más robusta para establecer comparaciones justas y equilibradas entre instituciones con características muy diversas.

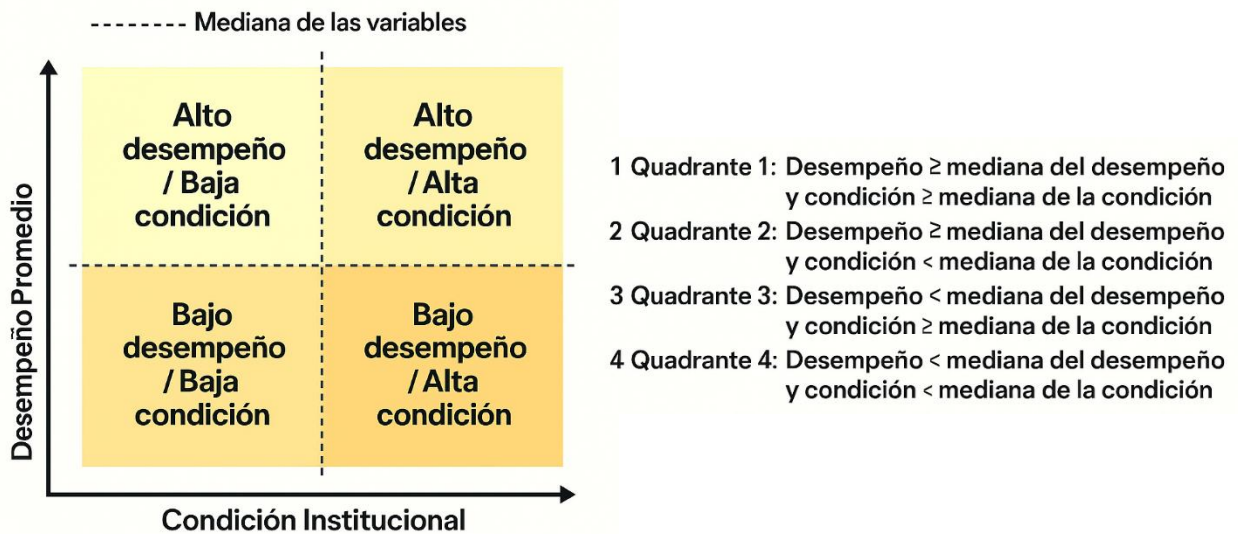
- **Mediana desempeño:** es la mediana (valor central) del desempeño de todas las instituciones.
- **Mediana condición:** es la mediana de la condición institucional de interés (por ejemplo, número de docentes, número de aulas, etc.).

Luego, se calculan los promedios de estas dos variables para la institución evaluada.

- **Promedio desempeño:** es el promedio del desempeño académico de interés.
- **Promedio condición:** es el promedio del valor de la condición institucional de interés.

A partir de estos valores se identifica en cuál de los cuatro cuadrantes se encuentra, tal como lo refiere la Figura 4:

Figura 4. Cuadrantes de relación.



Fuente: Elaboración propia

Posteriormente, se normalizan ambas variables en una escala de 0 a 1, para obtener una posición relativa dentro de su cuadrante, asignando el 50% del peso a cada variable. Esta posición se suma al valor base del cuadrante (100, 200, 300 o 400 según el caso), generando un resultado entre 100 y 500 que permite interpretar fácilmente el comportamiento de cada institución frente a la condición institucional analizada.

$$\text{Condición Normalizada} = \text{PromedioCondición} / \text{MáximoCondición}$$

$$\text{Desempeño Normalizado} = \text{PromedioDesempeño} / \text{MáximoDesempeño}$$

$$\text{Posición Relativa} = (\text{CondiciónNormalizada} * 50) + (\text{DesempeñoNormalizado} * 50)$$

$$\text{Resultado} = (\text{Número de Cuadrante} * 100) + \text{PosiciónRelativa}$$

Así, los resultados entre 100 y 200 indican alto desempeño con alta condición, entre 200 y 300 alto desempeño con baja condición, entre 300 y 400 bajo desempeño con alta condición, y entre 400 y 500 bajo desempeño con baja condición. Estos indicadores

facilitan el análisis comparativo por sectores y ayuda a identificar dónde hay eficiencia o necesidades de intervención según el contexto de cada institución (Ver Tabla 8).

Tabla 8. Indicadores de relación entre desempeño académico y las condiciones institucionales.

Indicador	Sigla	Descripción	Escala	Interpretación segmentada
Índice de cobertura de apoyo especial	ICAE	Mide si los apoyos a estudiantes con discapacidad cierran brechas.	100-500	100-200: Alto desempeño / alta cobertura 200-300: Alto desempeño / baja cobertura 300-400: Bajo desempeño / alta cobertura 400-500: Bajo desempeño / baja cobertura
Índice de desempeño académico docente	IDAD	Relación entre desempeño estudiantil y número de docentes.	100-500	100-200: Alto desempeño / más docentes 200-300: Alto desempeño / menos docentes 300-400: Bajo desempeño / más docentes 400-500: Bajo desempeño / menos docentes
Índice de estabilidad docente y rendimiento	IEDR	Impacto de la permanencia docente en el rendimiento.	100-500	100-200: Alto desempeño / alta estabilidad 200-300: Alto desempeño / baja estabilidad 300-400: Bajo desempeño / alta estabilidad 400-500: Bajo desempeño / baja estabilidad
Índice de disponibilidad de recursos y rendimiento académico	IDRRA	Influencia de recursos educativos en el aprendizaje.	100-500	100-200: Alto desempeño / muchos recursos 200-300: Alto desempeño / pocos recursos 300-400: Bajo desempeño / muchos recursos 400-500: Bajo desempeño / pocos recursos

Índice de desempeño según edad docente	IDED	Impacto de la edad docente en resultados académicos.	100-500	100-200: Alto desempeño / docentes mayores 200-300: Alto desempeño / docentes jóvenes 300-400: Bajo desempeño / docentes mayores 400-500: Bajo desempeño / docentes jóvenes
Indicador de resultados según el número de aulas	IRNA	Relación entre número de aulas y rendimiento.	100-500	100-200: Alto desempeño / más aulas 200-300: Alto desempeño / menos aulas 300-400: Bajo desempeño / más aulas 400-500: Bajo desempeño / menos aulas
Indicador de resultados según proporción directivos docentes	IRPDD	Influencia de directivos docentes en el rendimiento.	100-500	100-200: Alto desempeño / más directivos 200-300: Alto desempeño / menos directivos 300-400: Bajo desempeño / más directivos 400-500: Bajo desempeño / menos directivos
Indicador de resultados según número de estudiantes por aula	IRNEA	Efecto del tamaño del grupo en el desempeño.	100-500	100-200: Alto desempeño / más estudiantes 200-300: Alto desempeño / menos estudiantes 300-400: Bajo desempeño / más estudiantes 400-500: Bajo desempeño / menos estudiantes
Indicador de resultados según número de docentes en propiedad	IRNDP	Impacto de docentes con vínculo estable en rendimiento.	100-500	100-200: Alto desempeño / más docentes en propiedad 200-300: Alto desempeño / menos docentes 300-400: Bajo desempeño / más docentes 400-500: Bajo desempeño / menos docentes
Indicador de resultados según número de docentes en secundaria	IRNDS	Impacto de docentes de secundaria en resultados académicos.	100-500	100-200: Alto desempeño / más docentes de secundaria 200-300: Alto desempeño / menos docentes 300-400: Bajo desempeño / más

docentes400-500: Bajo
desempeño / menos
docentes

Fuente: Los indicadores fueron contruidos por el autor con base en criterios metodológicos de indicadores de segmentación tomados de referentes como Rodríguez Garcés et al., (2022), Simón Medina et al., (2022), De-La-Hoz et al., (2019) y Hinojosa et al., (2022).

7.6.2. Modelo de regresión

Se implementa un modelo de regresión múltiple para evaluar la relación entre el promedio del rendimiento académico en la prueba (variable respuesta) y diversas condiciones institucionales (variables predictoras). Este enfoque permite analizar la influencia de cada variable predictora a nivel institucional, facilitando la interpretación de su impacto en el desempeño académico.

Dentro del apoyo a los indicadores de relación la investigación emplea un modelo de predicción construido a partir de regresión múltiple y representado mediante visualizaciones interactivas. Este modelo permite identificar la influencia que ejercen diferentes condiciones institucionales sobre el desempeño académico en las pruebas. Para ello, se desarrollaron modelos entrenados con los datos institucionales disponibles, los cuales permiten estimar el rendimiento académico a partir de variables como: Número de docentes totales, promedio de tiempo de servicio en semanas, promedio de edad docentes, número de directivos docentes, número de estudiantes con alguna discapacidad o condición especial, número de recursos institucionales, número de estudiantes por aula, número de docentes propiedad, número de docentes secundaria, número de estudiantes por institución educativa, número de aulas.

8. Diseño de la arquitectura BI

8.1. Capa de fuente de datos

En esta capa recae principalmente la descripción de los archivos de origen, los cuales en esta investigación se conforman por tres archivos en formato CSV segmentados por semestre (2021-I, 2021-II, 2022-I y 2022-II) con información anonimizada de los resultados de la estrategia EPA. Adicionalmente, se incluyen dos pares de archivos en formato CSV con datos anonimizados de la planta docente y la matrícula académica.

8.2. Capa ETL (Extracción, transformación y carga)

En esta capa juega un papel importante el proceso de integración de datos, el cual se llevó a cabo con los tres procesos fundamentales del ETL.

8.2.1. Extracción

Se garantiza la obtención y consolidación de datos relevantes de la planta docente, matrícula, estrategia EPA, instituciones educativas y referencias geográficas. Se utiliza Python mediante *Jupyter Notebook* y bibliotecas como *os*, *pandas* y *numpy* para la manipulación eficiente de datos.

8.2.2. Transformación

Se depuran y transforman los datos en *Jupyter Notebook*. Se agregan columnas de referencia para años e instituciones educativas, se eliminan columnas innecesarias, se ajustan tipos de datos numéricos, se corrigen errores y se subsanan espacios nulos (NULL) y en blanco. Para los datos geográficos y las instituciones educativas, se realiza un proceso de jerarquización y combinación, consolidando la información en archivos únicos con los datos que caracterizan cada dimensión del modelo físico planteado en la siguiente sección.

8.2.3. Carga

En la carga del modelo BI es necesario hacer referencia del desarrollo del Almacén de Datos (DWH) para el cual inicialmente se debe establecer el modelo dimensional (modelo conceptual) con sus correspondientes hechos y dimensiones como lo muestra la Tabla 9. Posteriormente, se establece el modelo físico y se cargan los datos desde cada archivo independiente que contiene la estructura dimensional utilizando SQL Server Management Studio en su versión 9.1, asegurando una copia de seguridad (backup) y optimizando el proceso de carga desde una instancia local del servidor.

Tabla 9. Descripción modelo conceptual, Hechos y Dimensiones.

TIPO DE TABLAS	NOMBRE	CODIGO	PRECEDENTES
Dimensiones	Instituciones educativas	DIM_IIEE	Unificación de los archivos CVS con los códigos de la institución y sedes
Dimensiones	Tiempo	DIM_TIEMPO	Tabla creada directamente en el visualizador PBI
Dimensiones	Condiciones institucionales	DIM_CON_INST	Tabla diseñada con los datos de planta docente y matricula
Hechos	Resultados de la estrategia EPA	FACT_EPA	Tablas con los datos de la prueba EPA

Fuente: Elaboración propia

Una vez realizado el modelo dimensional (modelo conceptual) de los datos en sus correspondientes hechos y dimensiones, conformando así el DWH, se procede con la generación del modelo físico (Ver Figura 5).

Figura 5. Modelo Físico de datos.



Fuente: Elaboración propia.

8.2.3.1. Consideraciones para realizar el sistema de almacenamiento.

El diseño e implementación de un DWH es una etapa crítica en la construcción de una arquitectura de BI, ya que este sistema de almacenamiento centralizado permite consolidar, organizar y gestionar los datos de manera eficiente para su posterior análisis y visualización. A continuación, se detallan las consideraciones clave que se tuvieron en cuenta para el desarrollo del DWH en esta investigación:

- Para la construcción del DWH, es fundamental definir el modelo conceptual, el cual debe establecer la estructura lógica de los datos en términos de hechos y dimensiones.

Este modelo debe basarse en las necesidades específicas de la investigación, integrando los datos de la estrategia EPA y las condiciones institucionales. Las tablas principales del modelo conceptual deben incluir hechos, que contengan los datos medibles y cuantificables como los resultados de las pruebas EPA, y dimensiones, que proporcionen el contexto para los hechos, como instituciones educativas, tiempo y condiciones institucionales. Este enfoque dimensional permitirá organizar los datos de manera que faciliten consultas rápidas y análisis multidimensionales.

- Antes de cargar los datos en el DWH, se debe realizar un proceso exhaustivo de normalización y limpieza para garantizar la calidad y consistencia de la información. Esto implica la identificación y eliminación de registros repetidos en los archivos fuente, la corrección de valores erróneos o inconsistentes, como datos numéricos en campos de texto, la aplicación de estrategias para tratar valores faltantes, como la imputación de medias o la eliminación de registros incompletos, y la estandarización de los formatos de fechas, códigos y otros campos para garantizar coherencia entre los diferentes archivos.
- El proceso ETL es clave en la construcción del DWH y debe implementarse utilizando herramientas como Python con bibliotecas especializadas como *pandas* y *numpy*. La extracción debe consolidar los datos de múltiples fuentes, incluyendo archivos CVS (resultados EPA, matrícula, planta docente) y la descarga de datos externos desde la web (códigos DANE departamentales y municipales), además del uso de Jupyter Notebook para la manipulación de archivos. La transformación debe incluir la creación de columnas de referencia (años, instituciones educativas), la eliminación de columnas innecesarias y la agrupación de datos (medias, conteos). Finalmente, la carga de los datos transformados en SQL Server Management Studio debe garantizar una estructura optimizada para consultas y análisis.

- Para el almacenamiento de los datos, es recomendable seleccionar una plataforma robusta como SQL Server, debido a sus ventajas en términos de escalabilidad, integración con herramientas de BI como Power BI, seguridad en la gestión de usuarios y permisos, y facilidad para la realización de copias de seguridad y recuperación de datos en caso de fallos.
- La optimización del DWH es un aspecto clave para garantizar su rendimiento, por lo que se deben implementar estrategias como la indexación de tablas de hechos y dimensiones para acelerar las consultas, el particionamiento de tablas grandes en bloques más pequeños basados en criterios como el año o la región, y la normalización para reducir la redundancia de datos y mejorar la integridad.
- En términos de seguridad y control de acceso, será fundamental aplicar mecanismos de protección de datos adecuados. Se deberán definir roles y permisos en SQL Server para restringir el acceso a la información solo a usuarios autorizados. Además, en cumplimiento con normativas de privacidad, los datos sensibles de los estudiantes deberán estar anonimizados para evitar la identificación individual.
- Es importante considerar la integración del DWH con herramientas de análisis y visualización, asegurando que los datos estén estructurados en un formato compatible con Power BI, estableciendo conexiones directas entre SQL Server y Power BI para facilitar la actualización automática de los datos y optimizando las consultas para reducir los tiempos de carga en el tablero de control.

8.3. Capa de Análisis

En esta capa se realiza la aplicación de técnicas de análisis mencionadas en el capítulo 4.6, donde Power BI, manteniendo la estructura relacional diseñada

anteriormente, permite el desarrollo de objetos visuales y operaciones de análisis de variables.

8.3.1. Análisis exploratorio de datos (EDA) mediante indicadores claves

Para el caso, Power BI Desktop (aplicación de escritorio) manteniendo la estructura relacional diseñada anteriormente e importada directamente desde la instancia local en SQL Server, permite el desarrollo de objetos visuales dinámicos (gráficos, segmentaciones, cards, esquemas y medidores) que reflejan los distintos indicadores a analizar.

8.3.2. Análisis del modelo relacional

Para este análisis se debe aplicar una dinámica del *machine learning*, estructurada en una secuencia básica que se debe seguir, en primer lugar, se identifican las variables numéricas relevantes que representan factores institucionales y que pueden influir en el rendimiento académico de los estudiantes.

Se construye una matriz de correlaciones para evaluar la relación lineal entre cada variable predictora y el rendimiento académico, se emplea la función *pearsonr* de la librería *scipy.stats*. Los resultados pueden interpretarse de diferentes maneras: una correlación positiva (cercana a 1) indica que, a medida que la variable predictora aumenta, el rendimiento académico tiende a mejorar; una correlación negativa (cercana a -1) sugiere que un aumento en la variable predictora se asocia con un menor rendimiento académico; y una correlación cercana a 0 implica que no existe una relación lineal significativa.

Luego, se calcula el p-valor para determinar si la correlación obtenida es estadísticamente significativa, considerando una correlación como significativa cuando $p < 0.05$.

Posteriormente, se implementa un modelo de regresión lineal múltiple utilizando la clase *LinearRegression* de la librería *scikit-learn*. Este modelo evalúa la relación entre el rendimiento académico, como variable independiente, y las condiciones institucionales, como variables predictoras.

Para garantizar una evaluación robusta del modelo, los datos se dividen en dos conjuntos: un conjunto de entrenamiento (70%), que se utiliza para ajustar el modelo, y un conjunto de prueba (30%), que permite evaluar su capacidad predictiva en datos no vistos. El rendimiento del modelo se evalúa mediante dos métricas clave. El Error Cuadrático Medio (MSE) mide la diferencia entre los valores reales y los valores predichos, donde un valor menor indica una mayor precisión. Por su parte, el Coeficiente de Determinación (R^2) refleja qué porcentaje de la variabilidad en el rendimiento académico es explicado por el modelo, con valores cercanos a 1 indicando una mayor capacidad predictiva.

Los coeficientes de regresión obtenidos permiten interpretar la magnitud y dirección del impacto de cada variable predictora sobre el rendimiento académico. Un coeficiente positivo indica que un aumento en la variable predictora mejora el rendimiento, mientras que un coeficiente negativo sugiere lo contrario.

Finalmente, a partir del análisis relacional (tabla de relación), se generan visualizaciones para estimar los valores óptimos de las condiciones institucionales que favorecen un mejor rendimiento académico. Estas visualizaciones incluyen el manejo dinámico de las condiciones en formato numérico y el posible resultado de la prueba.

8.4. Capa de Distribución

Esta capa sintetiza el desarrollo de la arquitectura de BI mediante la implementación de un *tablero de control* interactivo y reportes detallados. El objetivo del *tablero de control* es presentar los resultados de la capa de análisis a través de informes interactivos que incluyen objetos visuales con el diseño de indicadores clave y análisis

predictivos. Para lograr una interfaz de usuario intuitiva y colaborativa, se han diseñado páginas individuales que agrupan y detallan los indicadores propuestos. El diseño del *tablero de control* se organiza en secciones principales que abarcan distintos tipos de indicadores y análisis dentro de los cuales encontramos:

8.4.1. Indicadores de desempeño académico

El cuadro de mando presenta una vista integral de los indicadores de desempeño académico en las pruebas EPA, permitiendo analizar los resultados de las instituciones educativas de Montería entre 2021 y 2022. El usuario encontrará ocho visualizaciones interactivas, acompañadas de filtros que permiten segmentar los datos por año, instrumento evaluado (matemáticas, inglés, lectura crítica, etc.), componente y competencia. Estas herramientas permiten explorar el rendimiento desde distintas perspectivas:

1. Treemap: Desempeño académico promedio por institución educativa. Representa jerárquicamente todas las instituciones según su rendimiento. El color indica el nivel de desempeño (escala cromática) y el tamaño refleja el promedio en la prueba.

2. Gráfico de columnas agrupadas: Desempeño académico por instrumento y competencia. Compara los resultados por área evaluada y sus competencias específicas mediante barras agrupadas, con una tabla complementaria que detalla cada valor.

3. Gráfico de barras: Desempeño académico promedio por grado. Presenta el rendimiento académico desglosado por nivel escolar (3° a 11°), acompañado de una tabla con el detalle de cada barra.

4. Gráfico de barras: Índice de desempeño por jornada escolar. Analiza los resultados según la jornada (mañana, tarde, única), con soporte de una tabla que revela el detalle de cada valor.

5. Gráfico circular: Desempeño promedio por zona. Compara visualmente el rendimiento entre zonas urbanas y rurales, con una escala cromática y el desempeño promedio.

6. Gráfico circular: Desempeño promedio por género. Representa comparativamente el rendimiento de estudiantes hombres y mujeres, mostrando porcentajes por género con una tabla de apoyo.

7. Gráfico circular: Desempeño por rango de rendimiento. Clasifica el desempeño en cuatro rangos cromáticos: muy bajo (<25%), bajo (25–50%), medio (50–75%) y alto (>75%), evidenciando la proporción de cada uno.

8. Gráfico de barras: Brecha de género por grado. Muestra las diferencias de rendimiento entre géneros por grado escolar. Las barras positivas indican ventaja masculina; las negativas, femenina. Se complementa con una tabla que cuantifica la diferencia y la favorabilidad.

8.4.2. Indicadores de recursos humanos

Este panel permite analizar los principales indicadores del cuerpo docente. El usuario encontrará seis visualizaciones interactivas, acompañadas de filtros para segmentar por año, instrumento evaluado, componente y competencia:

1. Gráfico circular: Porcentaje de docentes por género. Muestra visualmente la proporción de docentes hombres y mujeres, con una diferenciación cromática clara.

2. Gráfico de anillos: Proporción de docentes por tipo de contrato

Representa la proporción de docentes con contrato en propiedad y provisionales mediante un diseño dicromático.

3. Gráfico circular: Proporción de docentes por rango de edad. Presenta la distribución de docentes en rangos etarios: 18–30, 31–40, 41–50 y 51 años o más, usando colores distintivos.

4. Tarjeta (Card): Número promedio de estudiantes por docente. Muestra el número promedio (valor entero) de estudiantes asignados por docente, brindando una métrica clave de carga académica.

5. Gráfico de líneas: Desempeño académico docente por instrumento. Representa la concentración de docentes según su rendimiento promedio en cada instrumento, permitiendo observar tendencias en el eje tiempo-rendimiento.

6. Gráfico de columnas agrupadas: Proporción de docentes por área. Muestra visualmente la proporción de docentes por área académica, diferenciando cada categoría por color.

8.4.3. Indicadores de equidad

Este panel analiza cómo influyen las condiciones especiales de los estudiantes en su desempeño académico. Se presentan tres visualizaciones interactivas, filtrables por año, e institución:

1. Gráfico circular: Proporción de estudiantes según condición especial. Representa cromáticamente la distribución de estudiantes con condición: ninguna, SRPA (sistema de responsabilidad penal adolescente), discapacidad (física o cognitiva) y apoyo especial.

2. Treemap: Distribución de instituciones según condición especial. Muestra cómo se agrupan las instituciones según la cantidad y tipo de estudiantes con condiciones especiales.

3. Gráfico de columnas agrupadas: Desempeño según condición especial. Compara el promedio de resultados en la prueba entre los distintos grupos de estudiantes según su condición.

8.4.4. Indicadores socioeconómicos

Este panel permite explorar el impacto del estrato socioeconómico sobre el rendimiento académico. Incluye tres visualizaciones interactivas, filtrables por año, e institución:

1. Gráfico circular: Proporción de estudiantes por estrato. Muestra la distribución cromática de estudiantes según su estrato socioeconómico (1, 2, 3 y 4).

2. Gráfico de columnas agrupadas: Desempeño por estrato socioeconómico. Presenta el promedio de desempeño en la prueba, desagregado por estrato.

3. Treemap: Distribución institucional según estrato. Representa visualmente las instituciones educativas según la concentración de estudiantes por estrato.

8.4.5. Indicadores de relación entre condiciones institucionales y resultados de la prueba

En este módulo se explora cómo se relacionan diez variables institucionales con el rendimiento académico de los estudiantes. Para ello, se presentan diez visualizaciones interactivas, cada una con filtros por año e institución, que permiten un análisis dinámico y comparativo. Cada indicador está representado por un medidor dividido en cuatro segmentos cromáticos, con valores que van de 100 a 500. Estos segmentos permiten interpretar fácilmente el tipo de relación entre la condición institucional y los resultados académicos: desde alto desempeño con alta condición hasta bajo desempeño con baja condición. Además, se incluyen tarjetas informativas fijas e interactivas que ofrecen un

resumen claro del indicador y permiten profundizar en el detalle e interpretación de cada caso, así como las posibles recomendaciones.

8.4.6. Modelo predictivo de resultados

Este panel presenta la proyección del rendimiento académico con base en un modelo de regresión lineal múltiple previamente entrenado. El usuario podrá ajustar las condiciones institucionales según sus necesidades específicas:

Las visualizaciones muestran cómo distintas variables institucionales, agrupadas por categorías como cuerpo docente, infraestructura y otras, influyen en la predicción del desempeño. El diseño del *tablero de control* permite ajustar dinámicamente los parámetros y observar en tiempo real la variación esperada en los resultados.

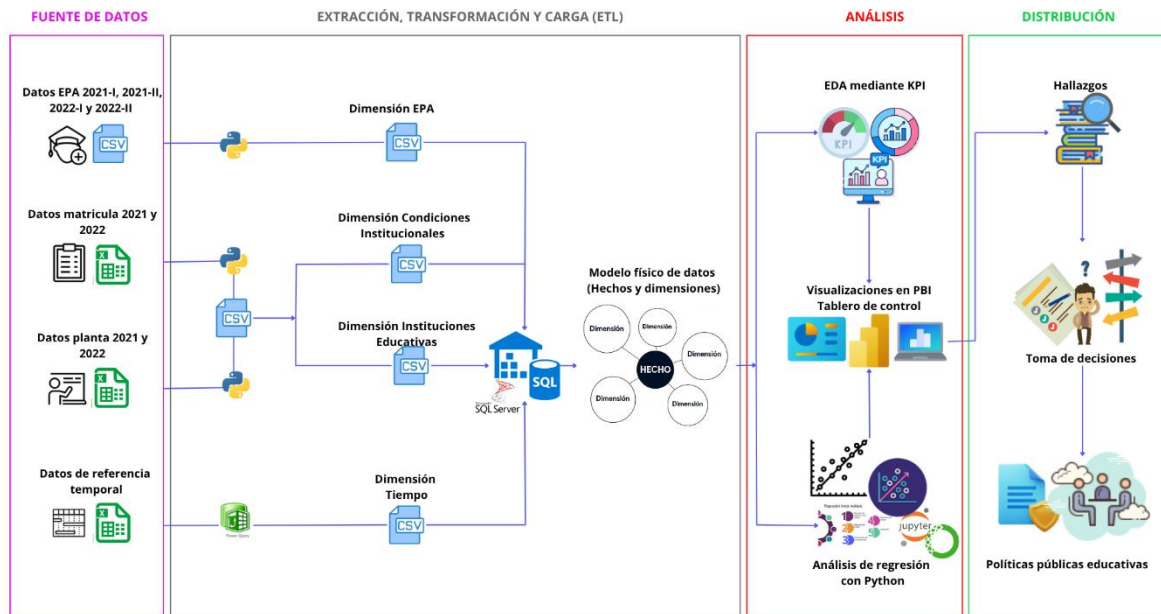
9. Trabajo de Campo

9.1. Procesamiento de los datos

En la Figura 6 se presenta el esquema de la arquitectura de inteligencia de negocios desarrollada para esta monografía, la cual estructura de manera integral el proceso metodológico implementado. Esta arquitectura no solo contempla los componentes técnicos como el almacén de datos (DWH), las tablas de hechos y dimensiones, y las herramientas de análisis como SQL Management Studio, Power BI y Python, sino que también incorpora el componente humano, clave en la comprensión e interpretación del contexto educativo.

En la capa final de esta arquitectura se encuentra el usuario decisor, encargado de interpretar los hallazgos y aplicarlos estratégicamente en el diseño de políticas públicas educativas. Además, los metadatos están presentes a lo largo de todas las capas del sistema, registrando aspectos esenciales como la fuente de los datos (SE Montería, ICFES, DANE), los formatos manejados (csv, txt, Excel), los años de referencia, las transformaciones realizadas y las versiones de las herramientas utilizadas. Esta trazabilidad no solo garantiza la integridad de los datos, sino que también fortalece la transparencia y reproducibilidad del análisis realizado.

Figura 6. Arquitectura BI aplicada en esta monografía.



Fuente: elaboración propia.

9.2. Análisis de indicadores

Dentro de los análisis de indicadores se destaca:

9.2.1. Indicadores de rendimiento

El análisis de los indicadores de desempeño académico en las pruebas EPA durante los años 2021 y 2022 revela una tendencia generalizada de bajo rendimiento. Más del 86% de las instituciones no alcanzan un 50% de desempeño en promedio (Ver Figura 7), lo que plantea un reto estructural para el sistema educativo de Montería. Esta situación se acentúa en el año 2022, donde el 93% de los estudiantes evaluados no superaron el umbral del 50%.

Figura 7. Desempeño académico promedio por rango.



Fuente: Elaboración propia. Tomado de la sección indicadores de rendimiento del tablero de control diseñado para este trabajo de grado.

La distribución del desempeño muestra diferencias territoriales y sociodemográficas importantes. Como lo evidencia la Figura 8, para el año 2021, las zonas urbanas presentaron un mejor desempeño (61,97%) que las rurales, a diferencia del año 2022, donde el rendimiento rural fue levemente superior (39,50%). Estas variaciones deben ser consideradas al momento de priorizar intervenciones territoriales.

Figura 8. Desempeño académico promedio según zona 2021 (izquierda) y 2022 (derecha).

Zona	Desempeño promedio en (%)	Zona	Desempeño promedio en (%)
Urbana	61,97	Rural	39,50
Rural	61,69	Urbana	39,38

Fuente: Elaboración propia. Tomado de la sección indicadores de rendimiento del tablero de control diseñado para este trabajo de grado.

En cuanto a la jornada académica, los datos muestran que no se presentan diferencias significativas en el desempeño académico general, ya que los promedios de resultados entre las jornadas mañana, tarde y única se mantienen relativamente estables (60% de rendimiento promedio en la prueba). Una excepción se observa en el año 2021 (Ver Figura 9), donde la jornada de la mañana reporta un leve aumento en el rendimiento respecto a las demás jornadas. Aunque esta diferencia no es contundente, podría sugerir la conveniencia de analizar más a fondo las condiciones pedagógicas, logísticas o contextuales que se presentan en dicha jornada, para identificar buenas prácticas replicables.

Figura 9. Desempeño académico promedio según jornada, año 2021.

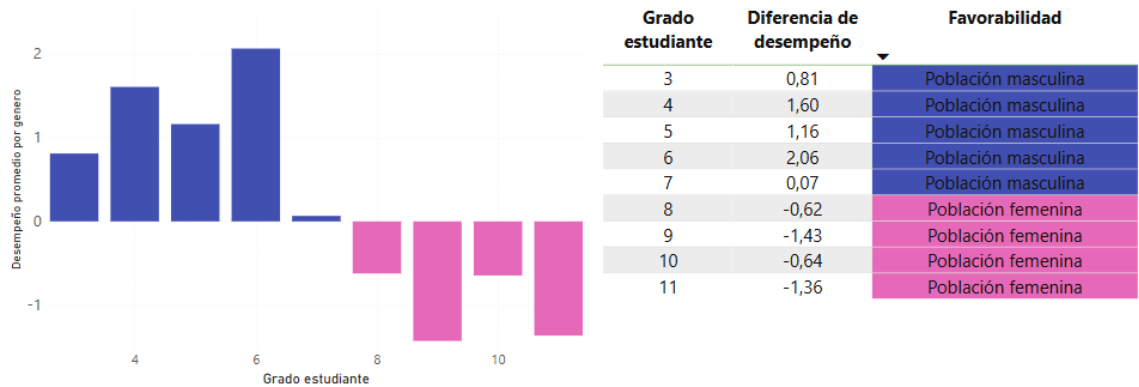
Promedio desempeño en la prueba	Jornada académica
62,23	MAÑANA
61,71	ÚNICA
61,39	TARDE

Fuente: Elaboración propia. Tomado de la sección indicadores de rendimiento del tablero de control diseñado para este trabajo de grado.

Respecto al género, los resultados son mixtos: en 2021, los hombres tuvieron mejor desempeño medio (62,03%) pero en 2022 las mujeres presentaron una ventaja leve de 0,65 puntos porcentuales. Sin embargo, la Figura 10, muestra la brecha de género por grado donde revela que las mujeres superan a los hombres en grados superiores, mientras que los hombres obtienen mejores resultados en los grados inferiores. Este hallazgo

sugiere la necesidad de implementar estrategias pedagógicas inclusivas y adaptativas, que respondan a las diversas formas en que el estudiantado aprende.

Figura 10. Brecha de género en los resultados de la prueba EPA por grado.



Fuente: Elaboración propia. Tomado de la sección indicadores de rendimiento del tablero de control diseñado para este trabajo de grado.

Los instrumentos con mejores resultados son ciencias naturales y educación ambiental, matemáticas e inglés como se evidencia en la Figura 11, aunque no superan un rendimiento medio del 60%. Esto indica un rezago transversal en competencias básicas, lo que refuerza la necesidad de reforzar estas áreas, especialmente en las instituciones con desempeño muy bajo.

Figura 11. Desempeño promedio de los distintos instrumentos (años conjuntos).

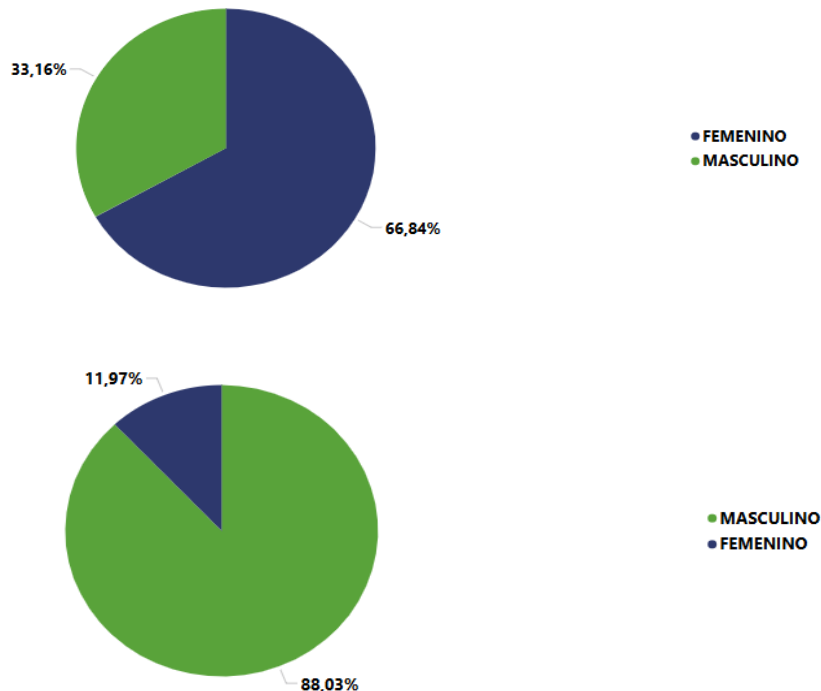
Instrumento	Desempeño promedio
Ciencias Naturales y Educación Ambiental	53,62
Matemáticas	48,55
Inglés	44,94
Competencias Comunicativas en Lenguaje:Lectura	43,25
Competencias Ciudadanas: Pensamiento Ciudadano	43,12
Ciencias Naturales	41,23
Sociales y Ciudadanas	40,78
Lectura Crítica	40,27

Fuente: Elaboración propia. Tomado de la sección indicadores de rendimiento del tablero de control diseñado para este trabajo de grado.

9.2.2. indicadores de recursos humanos

Los resultados del análisis de los indicadores de recursos humanos, señala inicialmente una distribución por género docente que muestra una variación significativa entre 2021 y 2022. Mientras que en 2021 el 66,84% de los docentes eran mujeres, en 2022 se observa una proporción mayoritaria de hombres, alcanzando el 88,03% (Ver Figura 12), lo que sugiere cambios estructurales en la vinculación docente que impactaron en la composición del equipo docente, lo que puede estar relacionado con procesos de rotación, provisiones o nombramientos.

Figura 12. Distribución docente por género año 2021 (superior) y 2022 (inferior).

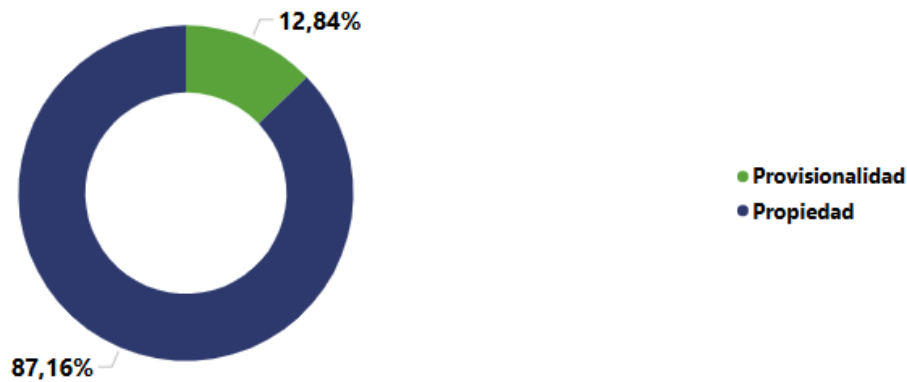


Fuente: Elaboración propia. Tomado de la sección indicadores de recursos humanos del tablero de control diseñado para este trabajo de grado.

Si bien no se identifican diferencias sustanciales en desempeño asociadas al género, este cambio en la estructura invita a fortalecer mecanismos de estabilidad y equidad en los procesos de selección y permanencia del talento humano.

La proporción de docentes vinculados en propiedad es mayoritaria, alcanzando el 87,16% del total (Ver Figura 13), lo cual ratifica a la estabilidad del personal docente, que constituye un aspecto positivo para la consolidación de planes académicos sostenibles y la implementación de estrategias pedagógicas a largo plazo.

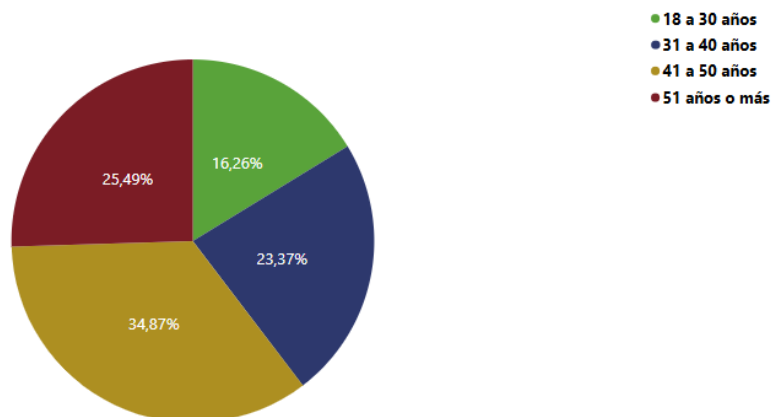
Figura 13. Proporción docente según contrato.



Fuente: Elaboración propia. Tomado de la sección indicadores de recursos humanos del tablero de control diseñado para este trabajo de grado.

En cuanto al rango etario, la mayoría de los docentes se ubican entre los 41 y 50 años (Ver Figura 14), representando el 34,87% del total, con un pico del 41,07% en 2021. Esto sugiere que el cuerpo docente cuenta con una experiencia acumulada significativa, lo cual representa una ventaja en términos de manejo de aula, conocimiento curricular y liderazgo pedagógico. Al mismo tiempo, este dato abre la posibilidad de proyectar estrategias de relevo generacional y formación de nuevos docentes, asegurando la continuidad de buenas prácticas y la transferencia de saberes institucionales.

Figura 14. Proporción docente según rango etario.



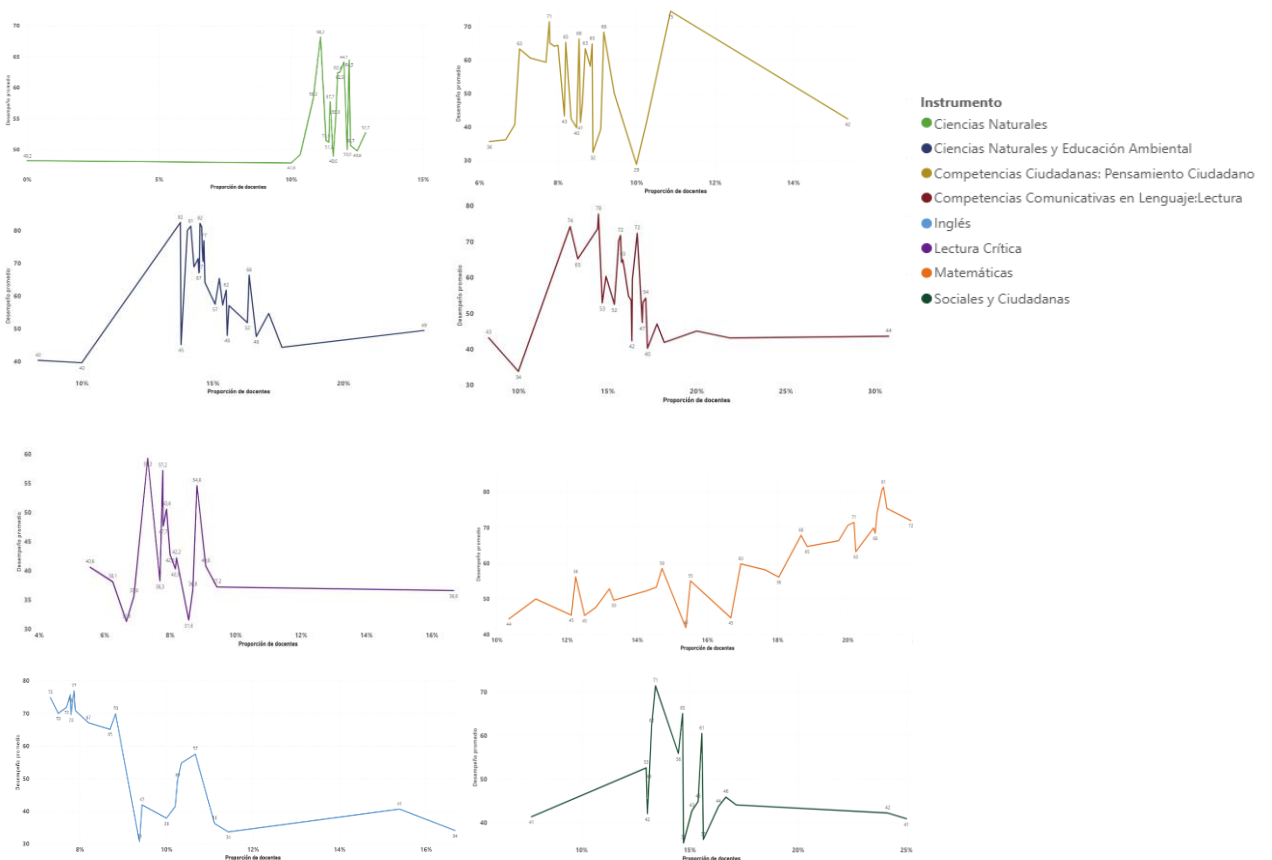
Fuente: Elaboración propia. Tomado de la sección indicadores de recursos humanos del tablero de control diseñado para este trabajo de grado.

El promedio de estudiantes por docente descendió de 13 en 2021 a 9 en 2022. Esta disminución sugiere una mejora en la distribución de la carga académica, lo que puede traducirse en una atención más personalizada al estudiante y un fortalecimiento de los procesos de enseñanza-aprendizaje. Sin embargo, para comprender completamente esta variación, es necesario revisar si obedece a una expansión de la planta docente, una disminución en la matrícula o una reorganización interna. En cualquiera de los casos, este indicador permite ajustar decisiones relacionadas con asignación de horas, tamaño de grupos y gestión del tiempo pedagógico.

El análisis del desempeño académico por instrumento revela diferencias importantes en la eficiencia del uso del recurso docente según el área evaluada. En 2021 (ver Figura 15), instrumentos como inglés, pensamiento ciudadano y lectura crítica lograron altos resultados con una baja proporción de docentes, lo que sugiere estrategias pedagógicas más eficaces o una mayor preparación del personal en estas áreas, sin

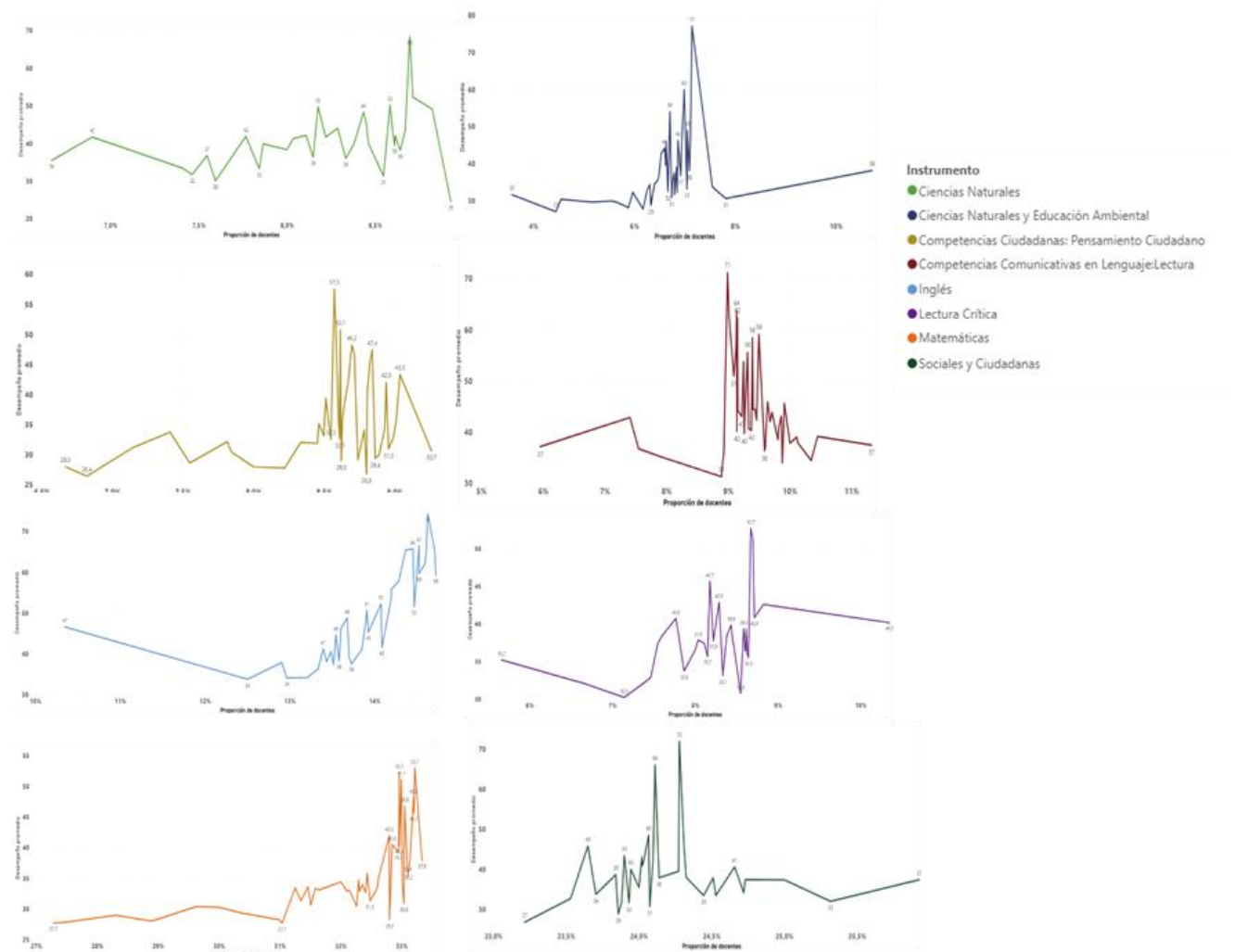
embargo, esta tendencia no se mantiene de manera constante en todas las instituciones ni en todos los periodos, lo cual sugiere la existencia de otros factores que pueden incidir en los resultados. Por el contrario, áreas como lectura, sociales y educación ambiental demandaron una mayor proporción de docentes con resultados más bajos, lo que plantea la necesidad de revisar las metodologías empleadas y ofrecer acompañamiento específico. En 2022 (ver Figura 16), se destaca matemáticas como el área que más docentes requirió (33%) con un desempeño máximo del 53%, lo que refuerza la urgencia de implementar planes de mejora y formación especializada en esta asignatura.

Figura 15. Proporción de docentes por área y desempeño, año 2021.



Fuente: Elaboración propia. Tomado de la sección indicadores de recursos humanos del tablero de control diseñado para este trabajo de grado.

Figura 16. Proporción de docentes por área y desempeño, año 2022.



Fuente Elaboración propia. Tomado de la sección indicadores de recursos humanos del tablero de control diseñado para este trabajo de grado.

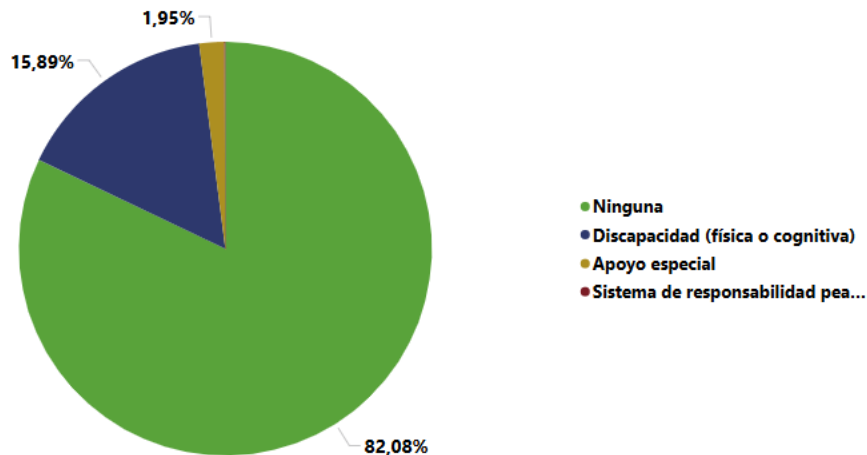
Finalmente, la proporción de docentes por área muestra que matemáticas concentra el mayor porcentaje de personal docente (30,44%), seguida de ciencias sociales (22,55%) y lenguaje (18,65%). No obstante, se destaca el área de inglés, que con solo un 13,21% de participación docente, presenta un desempeño sobresaliente en las pruebas. Esto indica, que se están aplicando estrategias efectivas que podrían ser replicadas o

adaptadas en otras áreas con mayor carga docente pero menor rendimiento. Esta observación también invita a revisar la pertinencia en la distribución del personal, asegurando que los recursos se alineen con los objetivos de mejora académica y no solo con la tradición organizacional.

9.2.3. indicador de equidad

Dentro de los hallazgos realizados en los diferentes indicadores se evidencia como la mayor proporción de estudiantes no presenta ninguna condición especial, representando el 82,08% del total. Le sigue el grupo de estudiantes con discapacidad física o cognitiva con un 15,89%, y en menor medida, aquellos que requieren apoyo especial, con un 1,95% tal como lo ilustra la Figura 17. Esto sugiere que, si bien la población con condiciones especiales representa una minoría, su presencia es significativa y debe ser tenida en cuenta en la planeación educativa, especialmente en términos de inclusión, accesibilidad curricular y ajustes razonables. La existencia de estos grupos implica la necesidad de fortalecer estrategias pedagógicas diferenciales que garanticen el derecho a una educación de calidad para todos los estudiantes.

Figura 17. Distribución de estudiantes en condición especial.



Fuente: Elaboración propia. Tomado de la sección indicadores de equidad del tablero de control diseñado para este trabajo de grado.

La distribución de instituciones según la cantidad y tipo de estudiantes con condiciones especiales evidencia una concentración desigual, donde algunas instituciones atienden a una proporción mayor de esta población (Ver Figura 18). Esto sugiere que hay centros educativos que enfrentan mayores retos en términos de inclusión y atención diferenciada, lo cual debe ser considerado al momento de asignar recursos, capacitar al personal docente y acompañar los procesos institucionales. Esta información es clave para construir políticas de equidad y apoyo focalizado que respondan de manera efectiva a la realidad de cada institución.

Figura 18. Distribución de instituciones según estudiantes en condición especial

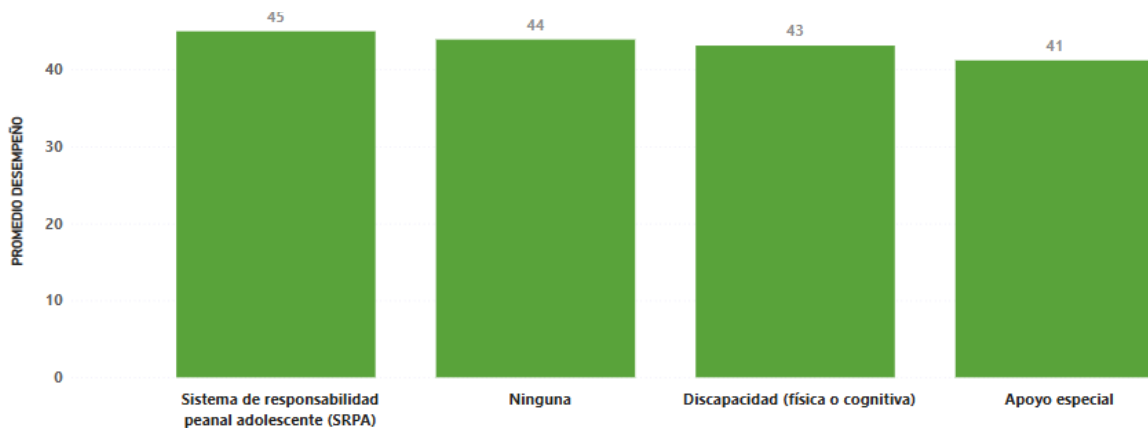


Fuente: Elaboración propia. Tomado de la sección indicadores de equidad del tablero de control diseñado para este trabajo de grado.

El análisis de desempeño académico según condición especial muestra que los estudiantes vinculados al Sistema de Responsabilidad Penal Adolescente (SRPA), aunque representan apenas el 0,08% de la población, obtienen los promedios más altos en las pruebas. Esto sugiere que la atención educativa en entornos con mayores niveles de intervención, como los del SRPA, podría estar favoreciendo una mayor concentración y seguimiento individualizado, aspectos que inciden positivamente en el rendimiento. Por otro lado, en las demás categorías de condición especial no se evidencian diferencias significativas, ya que las medias de desempeño se mantienen entre el 41% y el 44% en ambos años evaluados (Ver Figura 19). Sin embargo, es importante destacar que los estudiantes que requieren apoyo especial obtuvieron los promedios más bajos en los dos periodos analizados, lo que sugiere que este grupo necesita una intervención pedagógica más estructurada y específica. Esta situación puede estar asociada a una falta de recursos especializados o a limitaciones en la adecuación de estrategias didácticas, lo que

representa un llamado de atención para fortalecer la formación docente en inclusión y adaptar mejor las metodologías de enseñanza.

Figura 19. Desempeño académico promedio según tipo de condición especial.

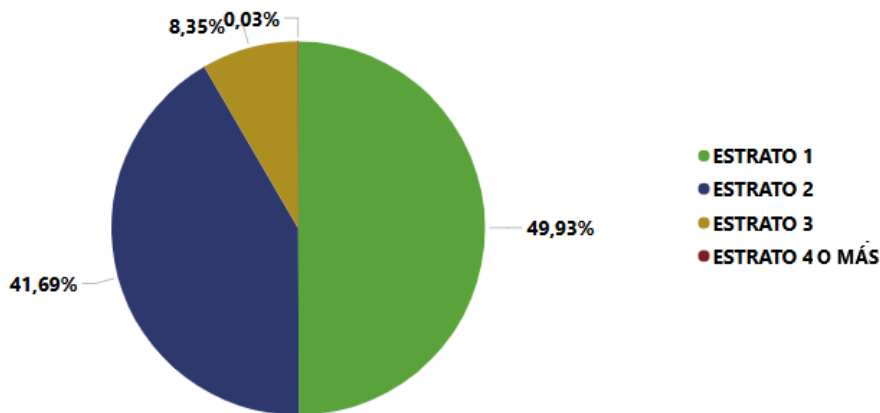


Fuente: Elaboración propia. Tomado de la sección indicadores de equidad del tablero de control diseñado para este trabajo de grado.

9.2.4. Indicador socioeconómico

El análisis de la proporción de estudiantes por estrato socioeconómico muestra que la mayoría de la población estudiantil pertenece a los estratos 1 y 2 (Ver Figura 20). El estrato 1 concentra el 50% de los estudiantes en ambos años analizados, seguido del estrato 2 con aproximadamente un 42%. Los estratos 3 y 4 tienen una representación mínima, siendo el estrato 4 o más prácticamente inexistente con un 0,03%. Indicando que el sistema educativo está atendiendo predominantemente a una población en condiciones de vulnerabilidad socioeconómica, lo que implica un reto importante para las instituciones educativas en cuanto al diseño e implementación de estrategias de acompañamiento integral que mitiguen los efectos de la desigualdad en el aprendizaje y los posibles problemas de deserción.

Figura 20. Proporción de estudiantes según estrato.

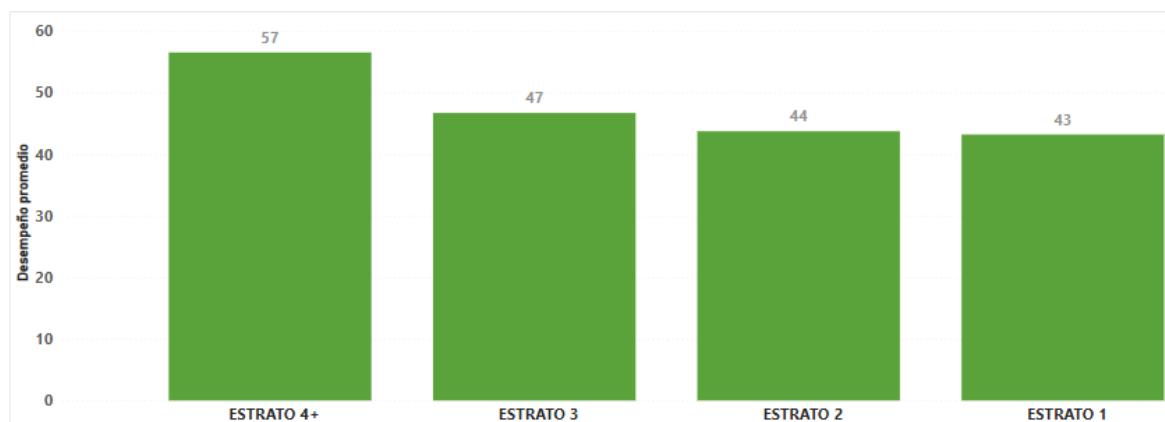


Fuente: Elaboración propia. Tomado de la sección indicadores socioeconómicos del tablero de control diseñado para este trabajo de grado.

En cuanto al desempeño académico por estrato socioeconómico, los resultados son sorprendentemente homogéneos. No se presentan diferencias significativas entre los promedios de desempeño de los estudiantes de estratos 1, 2 o 3 (Ver Figura 21). Esto sugiere que, a pesar de las limitaciones materiales que pueden enfrentar los estudiantes de menores recursos, el entorno escolar podría estar funcionando como un espacio de compensación, en donde la calidad del acompañamiento docente y las condiciones institucionales logran neutralizar, en parte, los efectos del contexto socioeconómico. No obstante, es importante señalar que esta aparente igualdad en los promedios no necesariamente indica que el desempeño académico sea alto o satisfactorio en términos absolutos. Más bien, invita a reflexionar sobre el techo de desempeño que podrían alcanzar los estudiantes si tuvieran condiciones aún más favorables. Por ejemplo, los estudiantes del estrato 4 presentan un desempeño académico ligeramente superior, aunque su representatividad en la muestra es limitada.

Es así como esta aparente igualdad en los promedios no debe hacer perder de vista la importancia de mantener políticas focalizadas que sigan apoyando a la población más vulnerable. En este sentido, las instituciones que deberían ser priorizadas por su alta concentración de estudiantes en situación de vulnerabilidad son: Normal Superior, José María Córdoba, Cecilia de Lleras, Antonio Nariño, Mercedes Abrego y La Pradera, ya que representan escenarios clave para el fortalecimiento de estrategias que promuevan la equidad educativa.

Figura 21. Desempeño académico promedio según estrato.



Fuente: Elaboración propia. Tomado de la sección indicadores socioeconómicos del tablero de control diseñado para este trabajo de grado.

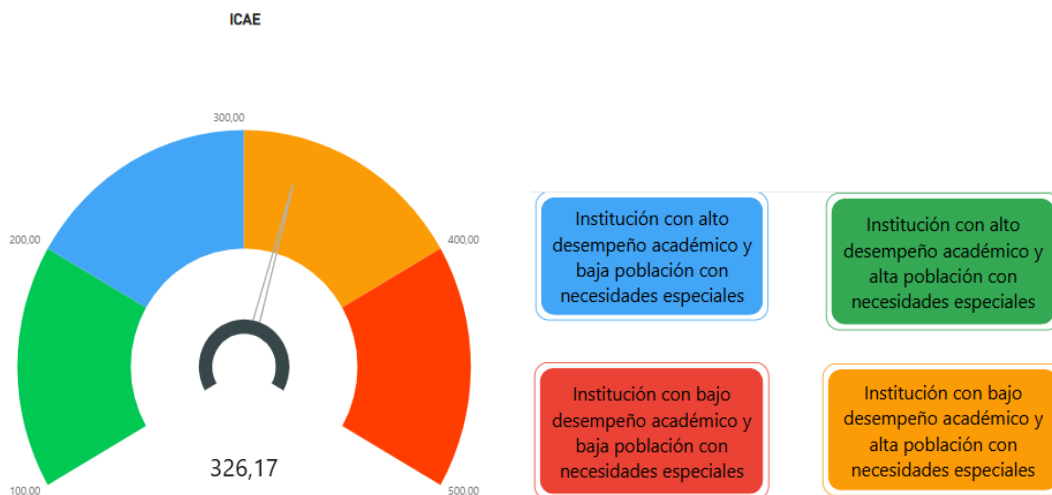
9.2.5. Indicadores de relación entre las condiciones institucionales y los resultados de desempeño de la estrategia EPA.

Los resultados globales de los diez indicadores analizados evidencian un patrón persistente de bajo desempeño académico en correlación con condiciones institucionales desfavorables o con alta presión estructural. Todos los indicadores cayeron en los cuadrantes de bajo resultado (por debajo del valor de la mediana del desempeño

académico), lo que confirma una problemática estructural sostenida que afecta a la mayoría de las instituciones educativas observadas.

El Índice de cobertura de apoyo especial (ICAE), con un valor de 326.17 en la categoría naranja (Ver Figura 22), sugiere que las instituciones con mayor población estudiantil con necesidades educativas especiales no logran traducir los apoyos en mejores resultados académicos. Este hallazgo pone en evidencia posibles falencias en la atención diferencial o una falta de recursos y formación especializada para docentes.

Figura 22. Indicador ICAE (Años conjuntos 2021 y 2022).



Fuente: Elaboración propia. Tomado de la sección indicadores de relación del tablero de control diseñado para este trabajo de grado.

En la misma línea, la Figura 23 muestra el índice de desempeño académico docente (IDAD) y el índice de estabilidad docente y rendimiento (IEDR) también se ubican en la categoría de bajo desempeño con alta carga docente (población docente elevada y antigüedad), lo que indica que la presencia o estabilidad de los docentes no está logrando un impacto positivo suficiente en el rendimiento estudiantil, tal vez por falta de

acompañamiento pedagógico, actualización profesional o condiciones de aula desfavorables.

Figura 23. Indicadores IDAD y IEDR (Años conjuntos 2021 y 2022).

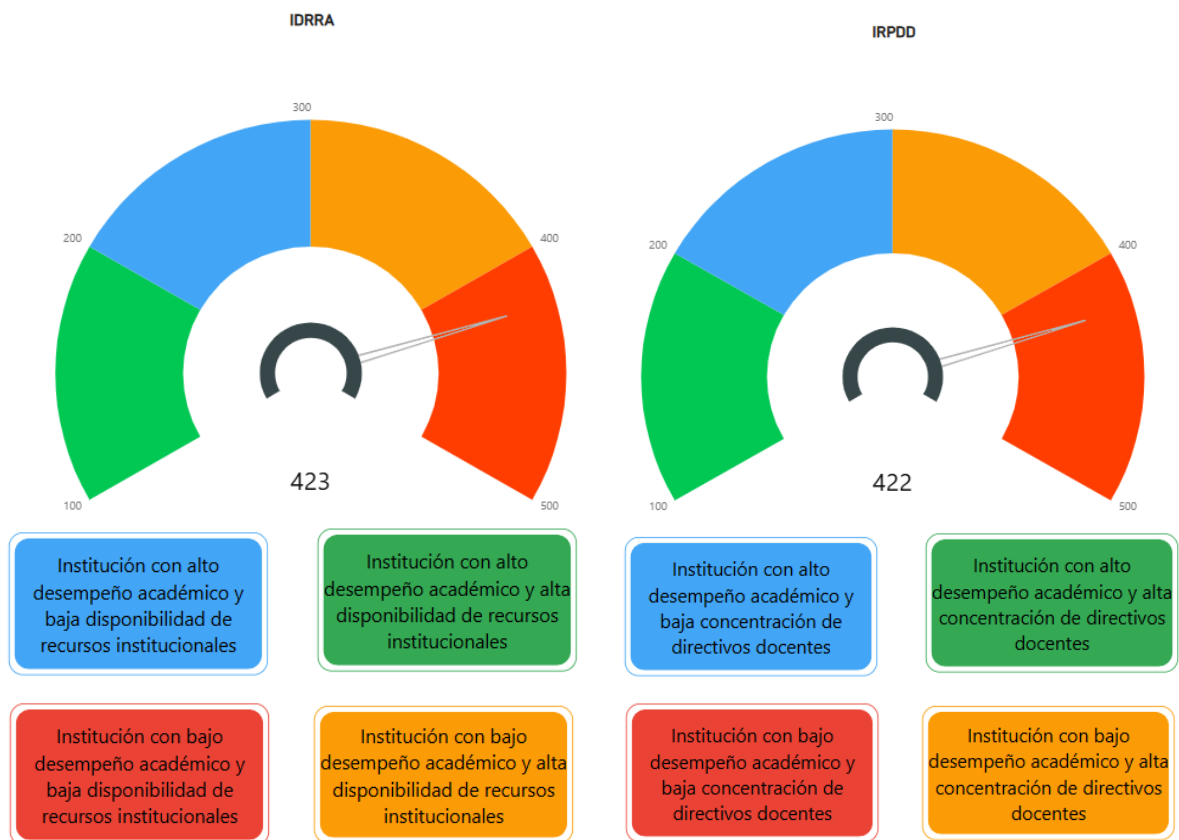


Fuente: Elaboración propia. Tomado de la sección indicadores de relación del tablero de control diseñado para este trabajo de grado.

Otros indicadores que se muestran en la Figura 24 y Figura 25, como el IDRRA (indicador que relaciona el desempeño con los recursos institucionales), IRNA (indicador que relaciona el desempeño con la cantidad de aulas), IRNEA (indicador que relaciona el

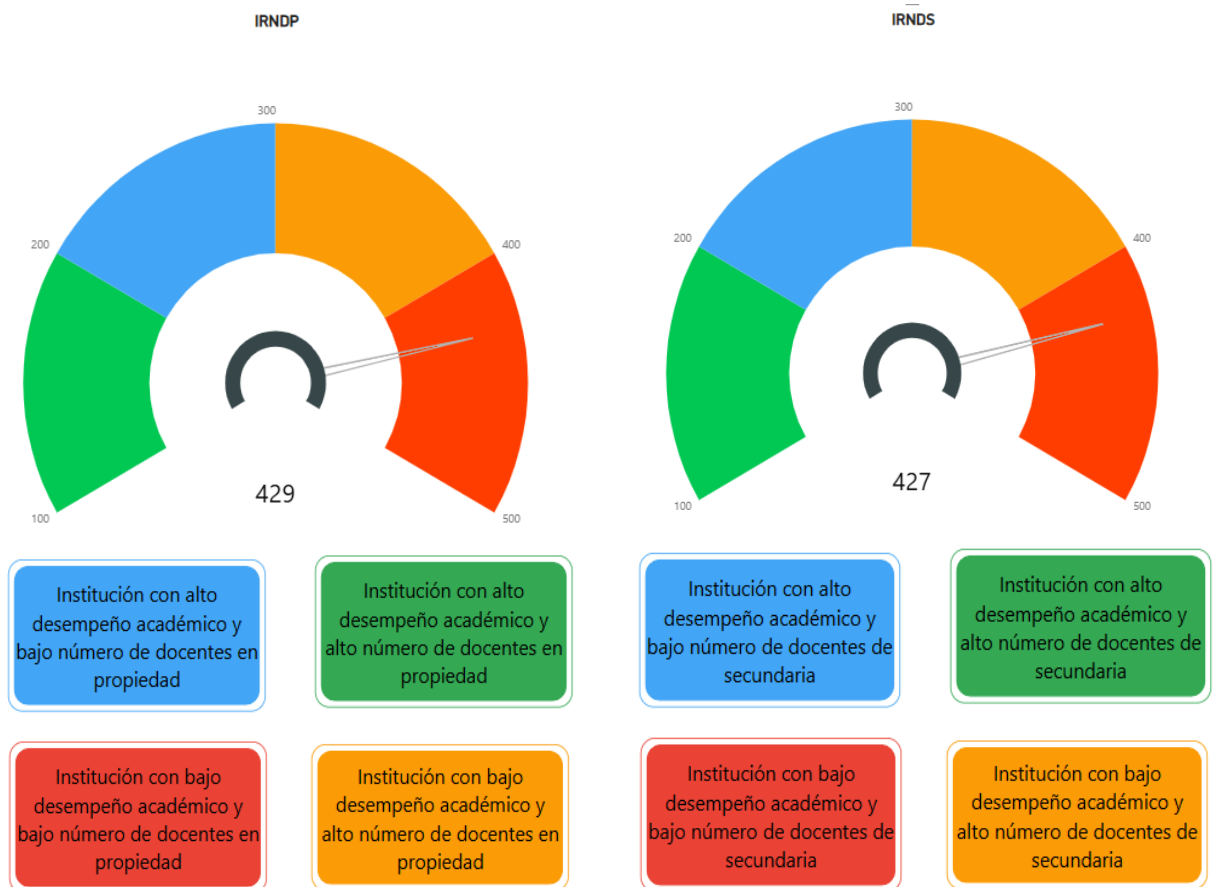
desempeño con el número de estudiantes por aula) y IRNDP/IRNDS (indicador que relaciona el desempeño con los docentes en propiedad y docentes de secundaria), todos en cuadrantes de bajo resultado, revelan una situación preocupante: no se observa un equilibrio entre los recursos físicos, humanos y organizacionales y los aprendizajes obtenidos. Particularmente crítico es el caso de los recursos físicos y humanos, donde incluso instituciones con alta dotación (como número de aulas o docentes) no alcanzan niveles de desempeño aceptables (por superiores a la mediana), lo cual sugiere que no es solo una cuestión de cantidad, sino de calidad, uso estratégico y eficiencia en la gestión de esos recursos.

Figura 24. Indicadores IDRRA y IRPDD (Años conjuntos 2021 y 2022).



Fuente: Elaboración propia. Tomado de la sección indicadores de relación del tablero de control diseñado para este trabajo de grado.

Figura 25. Indicadores IRNDP y IRNDS (Años conjuntos 2021 y 2022).

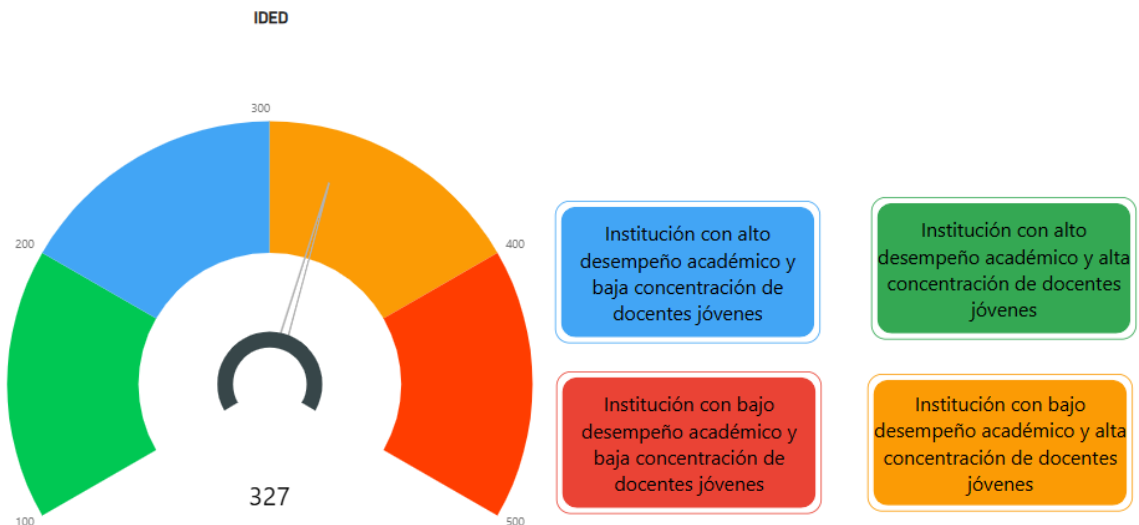


Fuente: Elaboración propia. Tomado de la sección indicadores de relación del tablero de control diseñado para este trabajo de grado.

Finalmente, la Figura 26 muestra el IDED (indicador que relaciona el desempeño con la edad de los docentes) evidencia que, aun cuando se cuenta con docentes jóvenes, lo que en teoría podría relacionarse con dinamismo o formación reciente, los resultados

siguen siendo bajos, lo que puede señalar la necesidad de fortalecer programas de inducción, mentoría y acompañamiento a nuevos docentes.

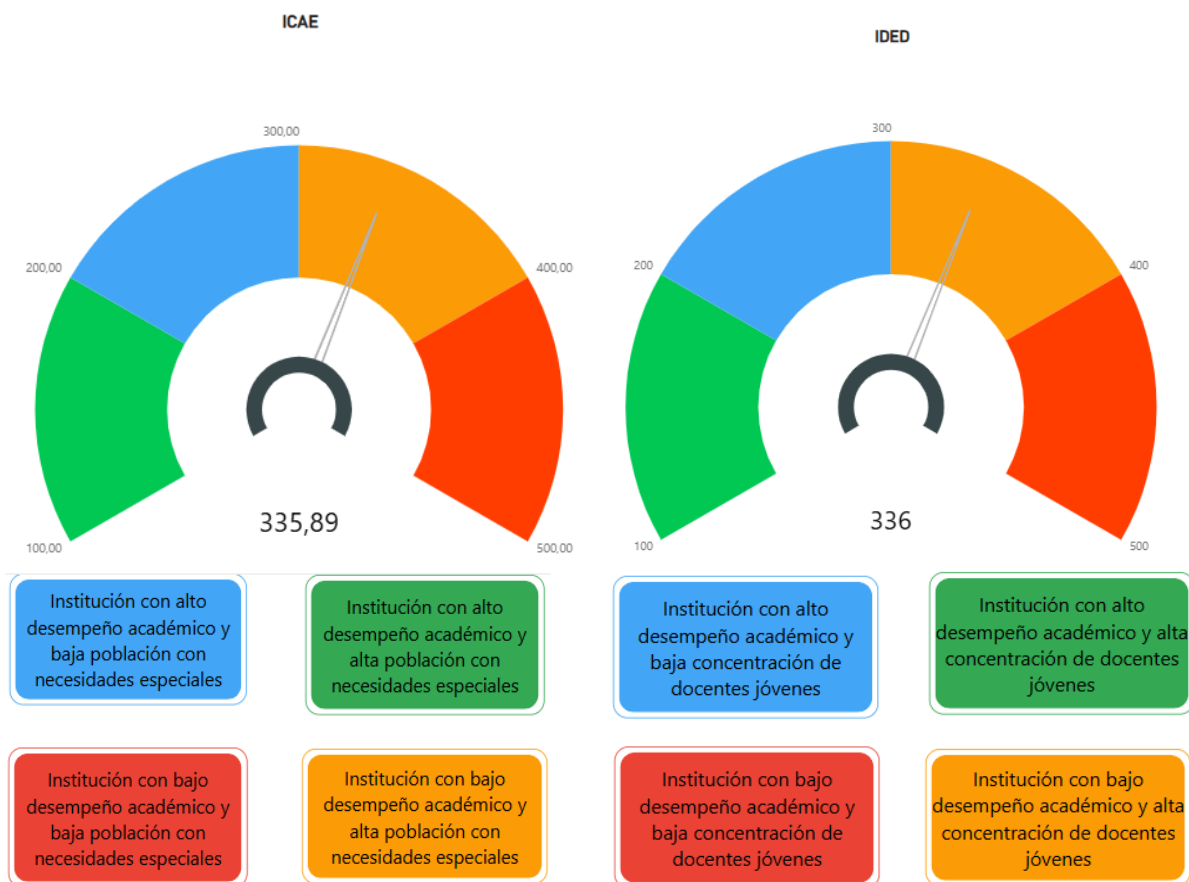
Figura 26. Indicador IDED (Años conjuntos 2021 y 2022).



Fuente: Elaboración propia. Tomado de la sección indicadores de relación del tablero de control diseñado para este trabajo de grado.

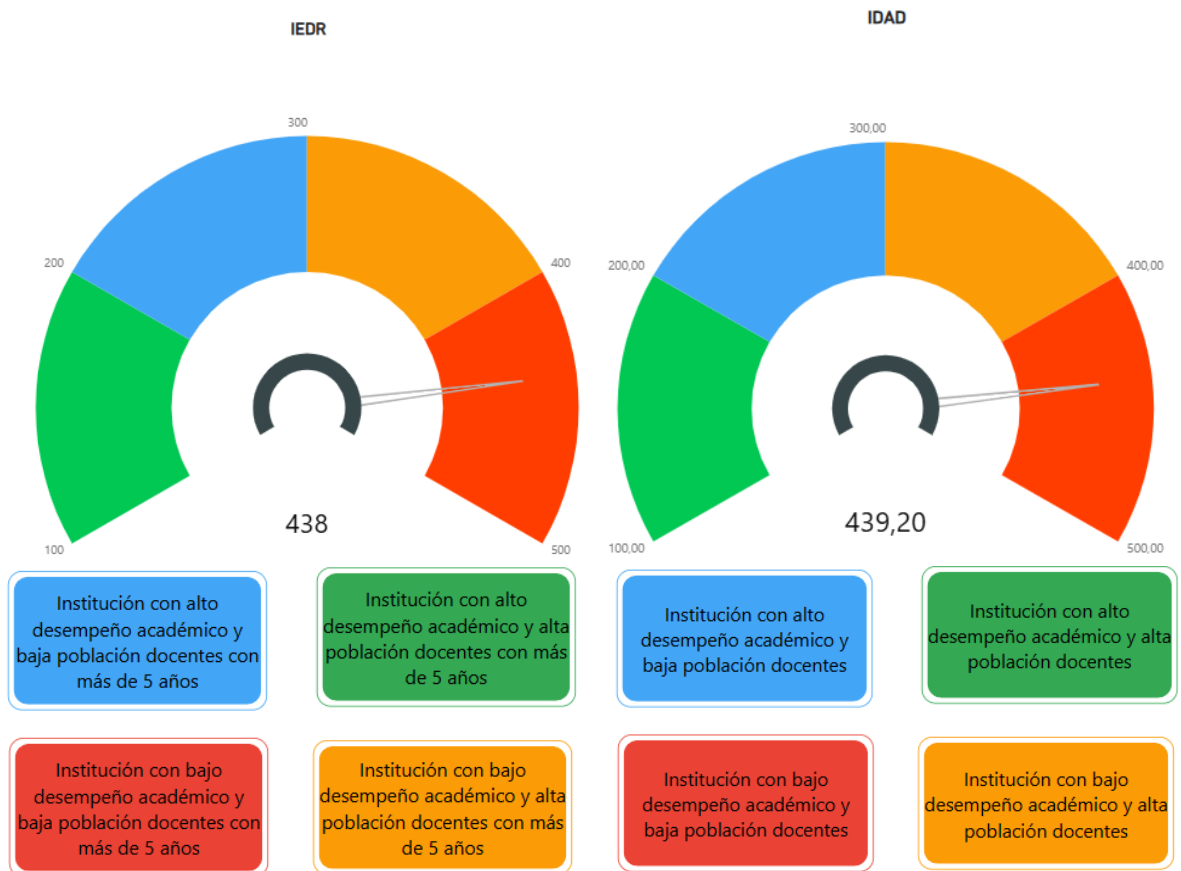
Asimismo, se evidenció durante el año 2021 que los indicadores institucionales reflejaban en su mayoría un escenario de bajo desempeño, a pesar de contar con altos niveles de cobertura o disponibilidad en diversos aspectos (Ver Figura 27). Por ejemplo, el Indicador de ICAE se encontraba en el cuadrante de 300 o categoría naranja, lo que representa un bajo desempeño, aún cuando la cobertura era amplia. Este mismo patrón se repetía en indicadores como el IDRRA, IDED, IRNA, IRPDD, IRNEA e IRNDS, los cuales también se encontraban en el cuadrante de 300. Incluso algunos como el IDAD, IERD e IRNDP, presentaban valores en el cuadrante 400 tal como lo refleja la Figura 28, que corresponden a bajo desempeño con baja disponibilidad, una combinación crítica que evidencia debilidad tanto en resultados como en recursos disponibles.

Figura 27. Indicadores ICAE e IDED, para el año 2021.



Fuente: Elaboración propia. Tomado de la sección indicadores de relación del tablero de control diseñado para este trabajo de grado.

Figura 28. Indicadores IEDR e IDAD, para el año 2021.

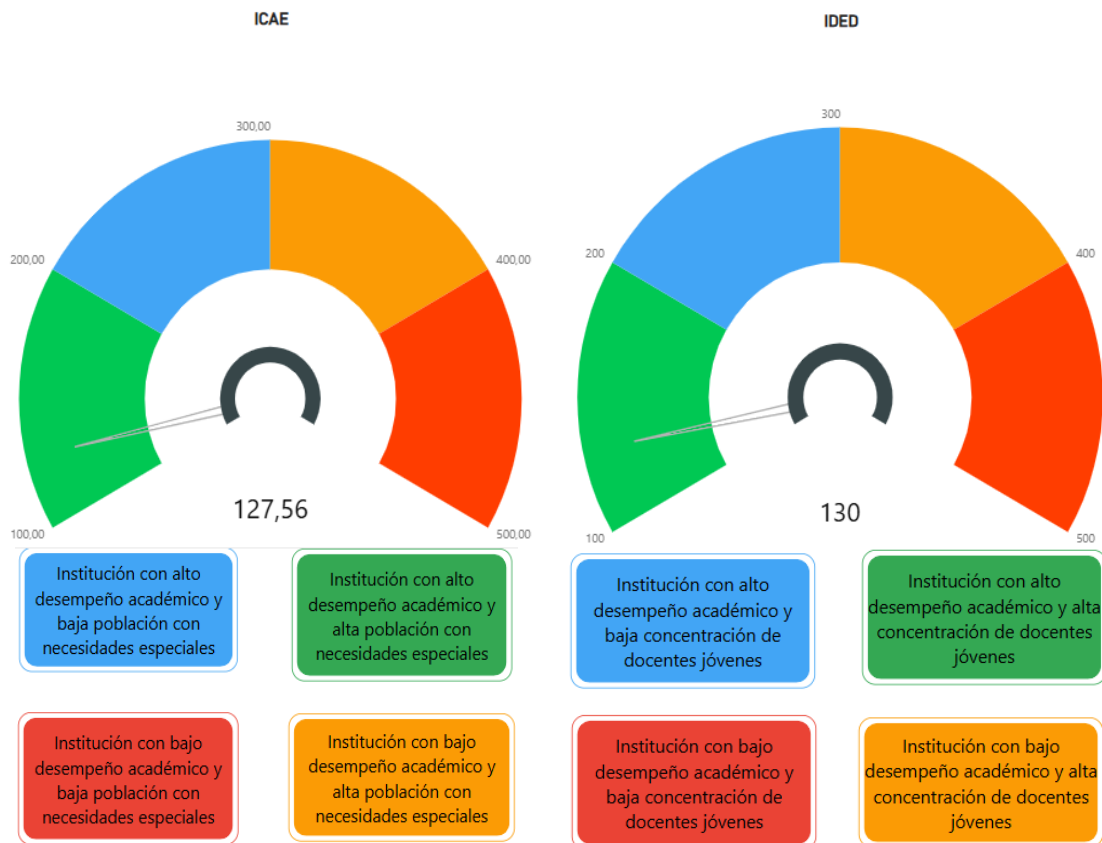


Fuente: Elaboración propia. Tomado de la sección indicadores de relación del tablero de control diseñado para este trabajo de grado.

Sin embargo, para el año 2022 se evidenció una mejora significativa en todos los indicadores evaluados, pues ninguno permaneció en el cuadrante de 300 o 400 (Ver Figura 29 y Figura 30), por el contrario, la mayoría descendió a los cuadrantes de 100 o 200, que corresponden a alto desempeño, ya sea con alta o baja disponibilidad o condición. El ICAE, por ejemplo, pasó a un indicador con valor 100 en 2022, lo que indica un alto desempeño acompañado de alta cobertura, reflejando una gestión una gestión más inclusiva y efectiva en términos de atención a esta población. Igualmente, el IEDR mejoró pasando de un indicador con valor de 400 en 2021 a 100 en 2022, lo que implica no solo una mejora en

la estabilidad del cuerpo docente, sino también una incidencia directa en los resultados institucionales. Casos similares se observan en el IRNEA, que pasó de un indicador con valor de 300 en 2021 a 100 para 2022, y en el IDED, que también descendió a valores en el cuadrante de 100, lo que sugiere una mayor efectividad en la gestión de grupos grandes y en la dirección escolar, respectivamente.

Figura 29. Indicadores ICAE e IDED, para el año 2022.



Fuente: Elaboración propia. Tomado de la sección indicadores de relación del tablero de control diseñado para este trabajo de grado.

Figura 30. Indicadores IEDR e IDAD, para el año 2022.

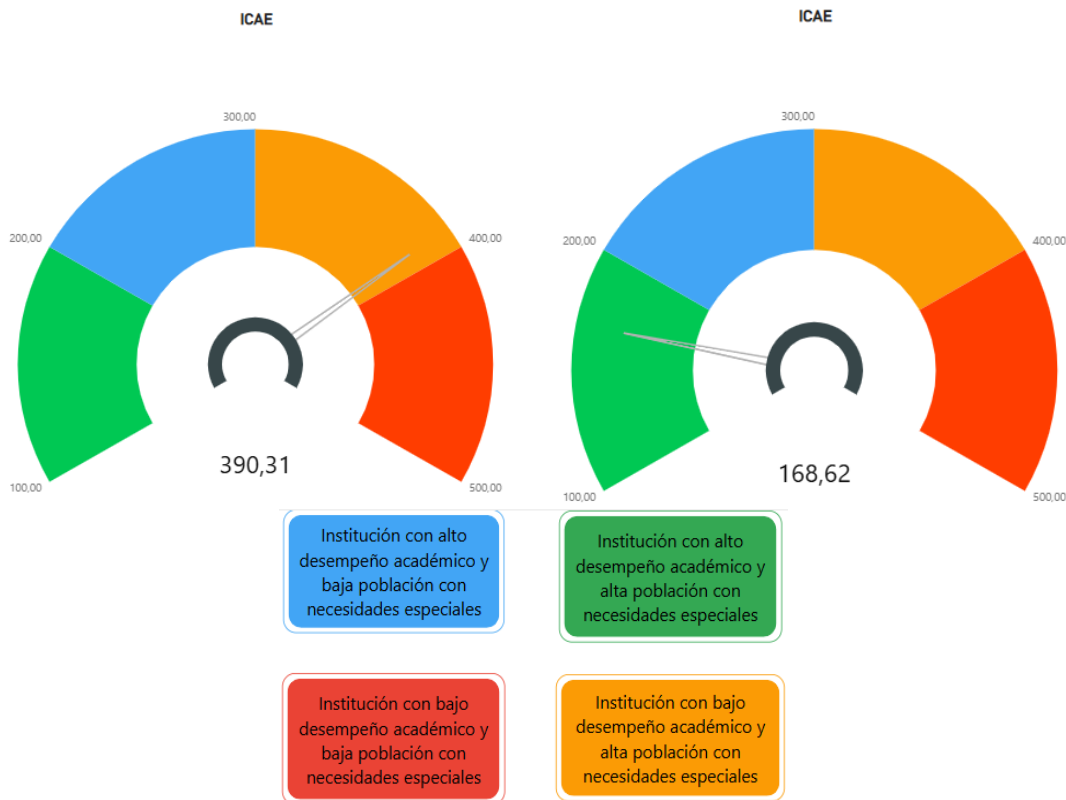


Fuente: Elaboración propia. Tomado de la sección indicadores de relación del tablero de control diseñado para este trabajo de grado.

El paso generalizado de la mayoría de los indicadores a los rangos de 100 o 200 sugiere un fortalecimiento de la eficiencia institucional. En particular, el hecho de que algunos indicadores se ubiquen en el cuadrante 200 (como IRNA, IRPDD, IDRRA, IDAD e IRNDP) indica que, aunque se cuenta con una menor disponibilidad o cobertura, los resultados logrados son altos, lo cual es un indicativo claro de eficiencia en la utilización de recursos limitados. Esta transformación refleja, posiblemente, mejoras en los procesos de planeación, gestión del talento humano, aprovechamiento de los recursos disponibles y fortalecimiento de los programas institucionales.

La visualización de los datos permite identificar de forma precisa las necesidades y avances individuales de cada institución educativa, como por ejemplo en el caso de la institución educativa Caño Viejo Palotal, que evidenció una mejora significativa en varios de sus indicadores clave entre los años 2021 y 2022. En el indicador ICAE, la institución pasó de un indicador de 390 en 2021, lo que la ubicaba en un rango de bajo desempeño con alta población en condiciones especiales, a un indicador de 168 en 2022 (Ver Figura 31), que representa un alto desempeño con adecuada cobertura de dicha población, lo que sugiere una optimización en el acceso y permanencia de los estudiantes en sus diferentes procesos educativos.

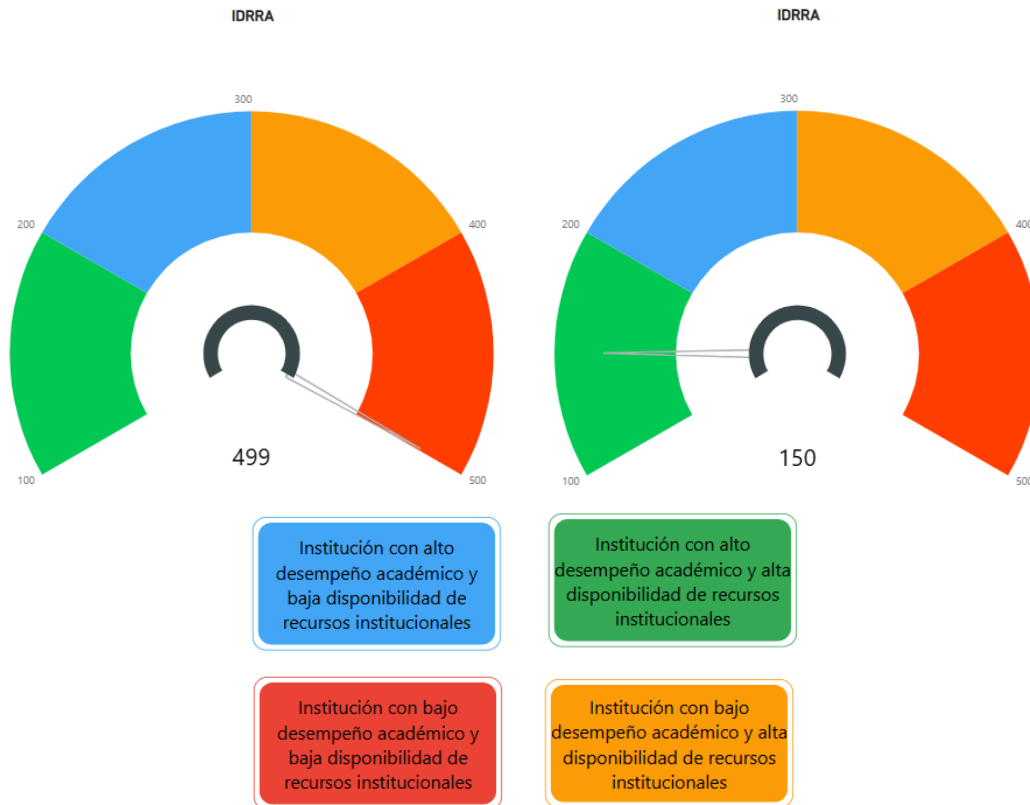
Figura 31. Comparativo indicador ICAE para la institución educativa Caño Viejo Palotal en los años 2021-2022 (izquierda-derecha).



Fuente: Elaboración propia. Tomado de la sección indicadores de relación del tablero de control diseñado para este trabajo de grado.

De manera similar, el indicador IDRRA mostró una mejora sustancial, al pasar de 498 para el año 2021 a 150 en el 2022 (Ver Figura 32), lo que implica que, a pesar de una posible limitación en recursos, se logró alcanzar un elevado desempeño institucional (resultados mayores a la mediana del desempeño académico), evidenciando una gestión eficiente y efectiva de los recursos existentes. El IRNA también descendió de 390 en 2021 a 246 para 2022, lo que ubica a la institución en un alto desempeño, posiblemente con una menor cobertura, lo cual sugiere que, incluso con menos beneficiarios, se lograron buenos resultados en términos de atención y calidad del servicio. Por último, el IRPDD mejoró drásticamente, pasando de un indicador en 498 en 2021 a 137 en 2022, lo que indica una mayor estabilidad en el liderazgo escolar, factor clave para la continuidad de procesos pedagógicos y administrativos.

Figura 32. Comparativo indicador IDRRA para la institución educativa Caño Viejo Palotal en los años 2021-2022 (izquierda-derecha).



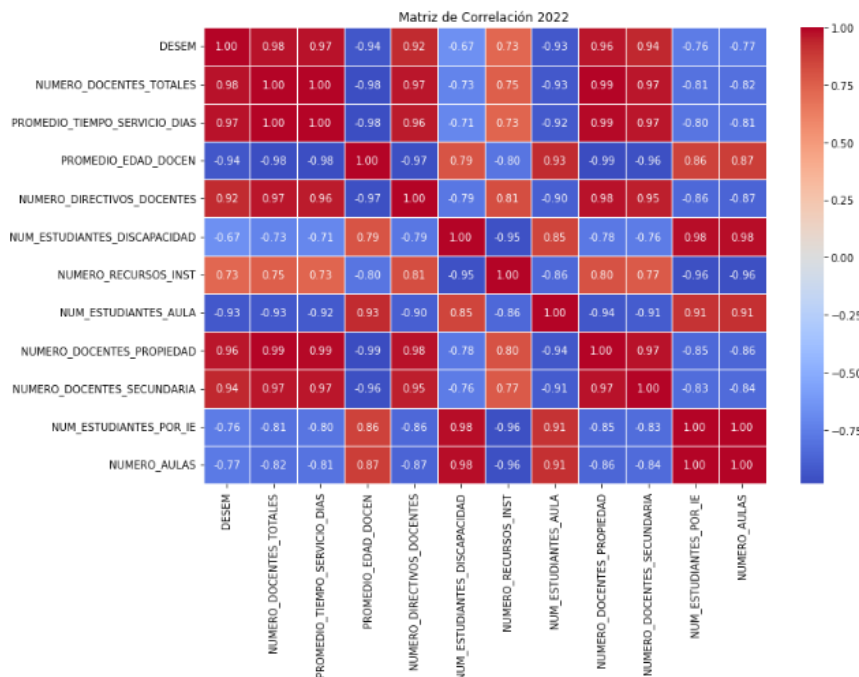
Fuente: Elaboración propia. Tomado de la sección indicadores de relación del tablero de control diseñado para este trabajo de grado.

Estos resultados sugieren que las instituciones han implementado estrategias efectivas de mejoramiento, probablemente relacionadas con una mejor gestión del talento humano, la reorganización de recursos y una planificación enfocada en el logro de resultados. La capacidad de transformar condiciones adversas en oportunidades de mejora refleja una dinámica institucional comprometida con la calidad educativa. Este tipo de análisis, además, sirve como referente para otras instituciones, al demostrar que es posible lograr un alto desempeño incluso en contextos de limitaciones estructurales, siempre que se cuente con liderazgo, planificación y voluntad de mejora continua.

9.3. Modelo predictivo

Dentro del análisis de resultados, apoyando la evaluación de la hipótesis de la investigación, se empleó correlación lineal de Pearson, donde se destaca la matriz de correlación la cual utilizó un modelo de correlación múltiple de las condiciones institucionales respecto al promedio de desempeño obtenido en la aplicación de la estrategia EPA de los distintos planteles educativos, dicha matriz corrobora el grado fuerte de correlación que oscilan entre 0.67 - 9.8 (Ver Figura 33), entre las distintas variables dependientes (condiciones institucionales) y la variable independiente (promedio desempeño prueba EPA por institución).

Figura 33. Matriz de correlación año.



Fuente: elaboración propia utilizando el ambiente Jupyter Notebook 6.4.8 .

Comprobada la colinealidad entre variables, se identifica que estas son significativamente estadísticas para explicar y predecir dentro del modelo ($p\text{-valor} < 0.05$), ya que el $p\text{-valor}$ de las variables oscila entre los valores de 0 y 0,01 y no sobrepasan el umbral de 0.05 de error estimado en estos casos.

Posteriormente, se implementó el modelo de regresión lineal con la librería *scikit-learn*, dividiendo los datos en un conjunto de entrenamiento (70%) y otro de prueba (30%) para garantizar una evaluación robusta, el rendimiento del modelo fue medido con el Error Cuadrático Medio (MSE) y el Coeficiente de Determinación (R^2), obteniendo resultados satisfactorios que respaldan su capacidad predictiva (Ver Tabla 10). No obstante, es importante aclarar que la muestra utilizada, aunque representativa, es limitada, por lo que los resultados podrían ser más robustos y generalizables con una mayor cantidad de datos.

Tabla 10. Métricas de rendimiento modelo de predicción.

Métricas de rendimiento del modelo
Error Cuadrático Medio (MSE): 7.23
Coeficiente de Determinación (R^2): 0.88

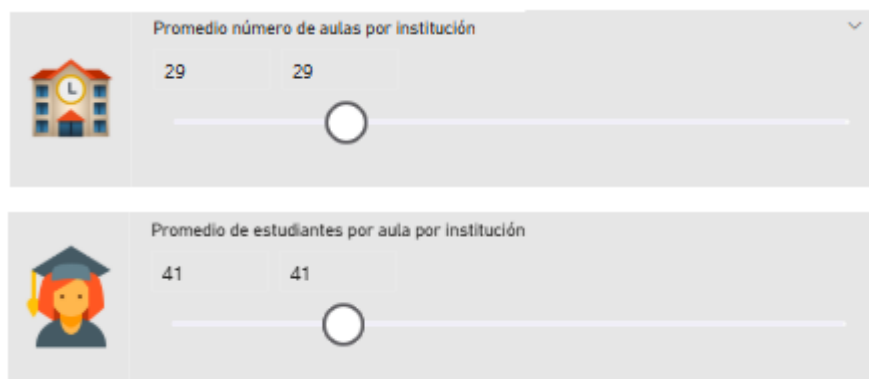
Fuente: elaboración propia utilizando el ambiente Jupyter Notebook 6.4.8.

Con base a los resultados del modelamiento de predicción se obtuvieron los siguientes resultados:

Las predicciones del desempeño académico promedio general de las instituciones educativas en Montería, que oscilan entre el 50% y el 65% (tomando como referencia una predicción promedio de 65,91% en las pruebas académicas), se derivan de un modelo de regresión lineal múltiple basado en técnicas de *machine learning*. Este modelo fue desarrollado con el propósito de apoyar a los administradores educativos en la

identificación de las condiciones institucionales que más inciden en los resultados académicos. El modelo indica que variables como el número promedio de aulas por institución y el número promedio de estudiantes por aula tienen un impacto muy alto en la predicción (Ver Figura 34), aceptando únicamente valores específicos (29 y 41 respectivamente), lo que revela una alta sensibilidad estructural en torno a estos factores. Asimismo, variables como el número promedio de recursos institucionales y el número de directivos docentes también muestran una influencia significativa, aunque en menor medida (Ver Figura 35). En contraste, otros elementos como el número total de docentes, el promedio de edad y tiempo de servicio del personal docente, así como la cantidad de docentes en propiedad o en secundaria, tienen un impacto bajo o nulo en la predicción. Esto sugiere que, dentro del modelo, no todas las condiciones tienen el mismo peso y que las decisiones orientadas a mejorar el rendimiento académico deben enfocarse principalmente en la gestión estratégica de la infraestructura, la densidad estudiantil por aula y los recursos institucionales disponibles. En esencia, este modelo no solo ofrece una predicción cuantitativa, sino que proporciona una guía práctica sobre las variables clave que deben ser intervenidas para lograr una mejora real en los resultados educativos.

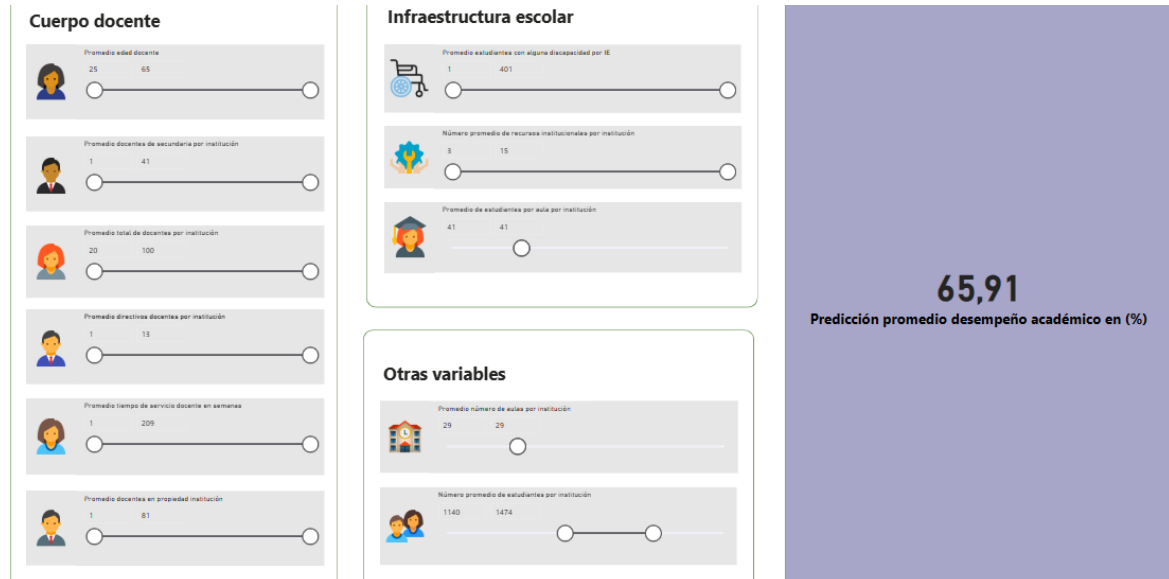
Figura 34. Condiciones institucionales sensibles dentro del modelo de predicción.



Fuente: Elaboración propia. Tomado de la sección predicción del tablero de control diseñado para este trabajo de grado.

El modelo no solo entrega una predicción cuantitativa del desempeño, sino que también ofrece una guía práctica sobre las variables estructurales prioritarias para intervenir con el fin de lograr mejoras sostenidas en la calidad educativa.

Figura 35. Predicción desempeño promedio del municipio de Montería.



Fuente: Elaboración propia. Tomado de la sección predicción del tablero de control diseñado para este trabajo de grado.

9.3.1. Propuesta de solución al problema planteado

La propuesta de investigación se centra en la implementación de una arquitectura de inteligencia de negocios diseñada para integrar los resultados de la estrategia “Evaluar para avanzar” con las condiciones institucionales de los centros educativos oficiales en Montería. Esta herramienta beneficia directamente a los usuarios finales, como directivos

docentes y secretarías de educación, al proporcionar datos consolidados, análisis predictivos y un tablero de control interactivo que facilite la toma de decisiones informadas. Por ejemplo, la arquitectura permite optimizar la distribución de recursos al analizar variables como el número de aulas, docentes y estudiantes, identificando instituciones con sobrepoblación o falta de personal, lo que permite redistribuir estudiantes o contratar más docentes de manera estratégica. Además, facilita la identificación de brechas en el desempeño académico al cruzar datos de resultados de pruebas con condiciones institucionales, como la edad de los docentes o la disponibilidad de recursos, lo que ayuda a priorizar intervenciones en áreas críticas.

La herramienta también permite focalizar intervenciones al identificar las variables con mayor impacto en el desempeño, como el número de docentes en propiedad o la estabilidad del personal, lo que asegura que las acciones tengan un efecto significativo. Asimismo, monitorea el impacto de las políticas implementadas mediante *un* tablero de control que muestran el progreso en tiempo real, lo que permite ajustar estrategias de manera ágil. La arquitectura promueve la transparencia y rendición de cuentas al generar reportes detallados que pueden ser compartidos con la comunidad educativa y los entes gubernamentales, demostrando el uso eficiente de los recursos y los avances logrados. Además, su diseño modular y escalable la hace adaptable a otros contextos educativos, ampliando su impacto más allá de Montería. Al ofrecer ideas claves, basados en evidencia, esta solución no solo mejora la gestión institucional, sino que también contribuye a visualizar el estado actual de las instituciones con miras a cerrar brechas y alcanzar los objetivos institucionales, transformando la manera en que se abordan los desafíos educativos en el municipio.

10. Directrices metodológicas para diseñar arquitecturas de inteligencia de negocios en el sector educativo en Colombia

En el sector educativo, el manejo de datos personales de estudiantes, docentes y personal administrativo es un tema de gran sensibilidad. Por ello, es fundamental garantizar que la arquitectura de BI cumpla con las normativas vigentes en Colombia, como la Ley 1581 de 2012 sobre protección de datos personales o el reglamento general de protección de datos (RGPD). Para ello, se deben establecer mecanismos que permitan obtener el consentimiento informado y explícito de los titulares de los datos antes de su recopilación y procesamiento. Además, es recomendable implementar técnicas de anonimización o pseudo anonimización de los datos, especialmente cuando se trabaja con información a nivel de estudiante, para preservar la privacidad. Asimismo, es crucial definir roles y permisos de acceso que aseguren que solo el personal autorizado pueda consultar y manipular los datos sensibles, minimizando así el riesgo de vulneraciones.

La estructura dimensional es un componente clave en el diseño de una arquitectura de BI, ya que define cómo se organizan y relacionan los datos para su análisis. En el contexto educativo, es importante establecer un nivel de granularidad adecuado, priorizando el detalle a nivel de estudiante como máximo y a nivel de institución educativa como mínimo. Esto permite un análisis profundo sin comprometer la privacidad de los individuos. El modelo dimensional debe incluir dos elementos principales: los hechos, que representan los datos medibles (como resultados de pruebas, asistencia o rendimiento académico), y las dimensiones, que son las categorías que permiten analizar estos hechos (como tiempo, estudiante, institución o área geográfica). Establecer relaciones claras entre estas dimensiones facilitará la generación de información valiosa para la toma de decisiones.

El diseño de la arquitectura de BI debe estar alineado con el contexto en el que se utilizará. Para ello, es necesario definir claramente los objetivos de la arquitectura, ya sea para mejorar la toma de decisiones institucionales, monitorear el rendimiento académico o identificar tendencias en el sector educativo. Además, es fundamental identificar a los usuarios finales del sistema, quienes pueden ser directivos de instituciones educativas, entidades gubernamentales o investigadores. Una vez definidos los objetivos y el público objetivo, se deben establecer los indicadores clave de desempeño (KPIs) relevantes para el sector educativo, como tasas de deserción, resultados de pruebas estandarizadas o eficiencia administrativa. Estos KPIs servirán como base para el análisis y la visualización de datos.

El diseño de la arquitectura de BI debe basarse en los resultados de pruebas piloto o análisis preliminares, lo que permitirá identificar las dimensiones y hechos más relevantes. Es esencial que la arquitectura permita la integración de datos provenientes de diversas fuentes, como sistemas de gestión académica, encuestas o pruebas estandarizadas, para garantizar un análisis integral. La base de datos debe estar diseñada para soportar el modelo dimensional y permitir la actualización y consulta eficiente de los datos. Además, se recomienda considerar la inclusión de hechos adicionales que enriquezcan el análisis, como condiciones socioeconómicas de los estudiantes o indicadores institucionales, siempre que estos estén relacionados con el mayor grado de granularidad posible.

La selección de herramientas tecnológicas es un aspecto crítico en el diseño de la arquitectura de BI. Es importante elegir herramientas que sean compatibles con el procesamiento de grandes volúmenes de datos y que permitan la visualización intuitiva de los resultados. Algunas opciones recomendadas incluyen Power BI para la creación del

tablero de control interactivo, Python (con Conda) para el análisis avanzado de datos y la implementación de modelos predictivos, y SQL Server para la gestión y consulta de bases de datos relacionales. Además, es fundamental asegurar que estas herramientas puedan integrarse entre sí y con los sistemas existentes en las instituciones educativas, garantizando así la interoperabilidad del sistema.

La escalabilidad es un aspecto esencial en el diseño de la arquitectura de BI, ya que permite adaptar el sistema a futuras necesidades. En términos de tiempo, la arquitectura debe ser capaz de incorporar nuevos datos o actualizar dimensiones sin requerir cambios estructurales significativos. En cuanto a los datos, debe permitir la adición de nuevos hechos, como resultados de nuevas pruebas, y dimensiones, como indicadores de calidad educativa. Además, la arquitectura debe ser lo suficientemente flexible para adaptarse a cambios en las políticas educativas o en los requerimientos de los usuarios, garantizando así su utilidad a largo plazo.

La implementación de una arquitectura de BI en el sector educativo colombiano tiene el potencial de generar un impacto significativo en la mejora de la calidad educativa. Sin embargo, es importante reconocer los retos y limitaciones asociados, como la resistencia al cambio, la falta de capacitación en el uso de herramientas de BI o la disponibilidad de datos de calidad. Para superar estos desafíos, se recomienda implementar acciones concretas, como la capacitación del personal y la estandarización de procesos de recolección de datos. En conclusión, las directrices presentadas en este capítulo proporcionan un marco metodológico robusto para el diseño de arquitecturas de BI en el sector educativo, contribuyendo así a la toma de decisiones basada en datos y al fortalecimiento del sistema educativo en Colombia.

11. Conclusiones, recomendaciones y trabajos futuros

11.1. Conclusiones

Se destaca la necesidad de desarrollar referentes teóricos que respondan a las particularidades del sector educativo, como la variabilidad de las condiciones institucionales que influyen en el rendimiento académico. La falta de metodologías adaptadas a estas condiciones puede limitar la eficacia de las arquitecturas de BI aplicadas al contexto escolar. Con esto se cumple el objetivo de establecer referentes teóricos para el diseño de una arquitectura de inteligencia de negocios adaptada a la realidad educativa.

A pesar de la disponibilidad de datos en el sector público, se identificó una explotación limitada de la información, centrada principalmente en evaluaciones básicas. Este estudio aplicó procesos de extracción, transformación y carga de datos que permitieron contrastar los resultados de la estrategia “Evaluar para Avanzar” con variables institucionales, mostrando el valor de una integración más profunda entre fuentes. Esto confirma el cumplimiento del objetivo de caracterizar las fuentes de datos y aplicar procesos ETL a los datos relevantes.

Se diseñó e implementó un modelo relacional de almacenamiento de datos basado en hechos y dimensiones, que permitió integrar los resultados de las pruebas EPA con las condiciones institucionales de los centros educativos oficiales de Montería. Esta arquitectura permitió cumplir el objetivo de configurar un sistema de almacenamiento eficiente y contextualizado.

La integración de herramientas como Python, Power BI y SQL Management permitió transformar y analizar los datos de forma flexible, generando visualizaciones que revelan patrones relevantes en el rendimiento académico en función de diferentes

variables institucionales. Esto da cumplimiento al objetivo de generar visualizaciones que articulen los resultados de evaluación con las condiciones de las instituciones.

La experiencia desarrollada en este estudio permite proponer directrices metodológicas para diseñar arquitecturas de inteligencia de negocios adaptadas al sector educativo colombiano, teniendo en cuenta tanto los aspectos técnicos como contextuales. De esta manera, se cumple el objetivo de establecer lineamientos que sirvan como guía para futuras implementaciones en entornos educativos.

11.2. Recomendaciones

Considerando la falta de enfoques específicos en la literatura actual para la integración de estrategias de evaluación, como la EPA, con las condiciones institucionales, se recomienda desarrollar arquitecturas de BI que se adapten específicamente al entorno educativo. Esto implica la creación de marcos teóricos que consideren las particularidades del sector, así como la apropiación de las consideraciones metodológicas presentadas en el capítulo anterior.

Asimismo, ante la creciente disponibilidad de datos, especialmente en el sector público, y la falta de explotación integral, se recomienda implementar estrategias que trasciendan la evaluación de aspectos básicos. Se sugiere aprovechar completamente el potencial de contrastar datos provenientes de diferentes áreas o sectores para obtener una visión más completa y rica para la toma de decisiones.

Se plantea la escalabilidad de la arquitectura a niveles territoriales, regionales y nacionales. Además, se recomienda ampliar las dimensiones analizadas; por ejemplo, incorporando datos geográficos para interpretar gráficamente las incidencias de las variables. También se propone la posibilidad de incluir datos más granulares en la tabla de hechos relacionada con la estrategia EPA.

Enfatizando la necesidad, se destaca la importancia de integrar de manera fluida la arquitectura con herramientas analíticas como Python notebook, Power BI y SQL Management. Esta integración puede potenciar el análisis y la visualización de datos, aprovechando las capacidades complementarias de estas herramientas.

Dada la evolución constante en el ámbito tecnológico, se recomienda establecer programas de formación continua para los profesionales involucrados en el manejo de la arquitectura de BI. Esto garantizará la adaptación a nuevas tecnologías y metodologías, manteniendo a los equipos actualizados y preparados para los desafíos tecnológicos emergentes.

11.3. Trabajos futuros

Para futuros trabajos, se mantiene la posibilidad de escalar la arquitectura a niveles territorial, regional y nacional. Las dimensiones analizadas permiten su expansión tanto lógica como física. Por ejemplo, se puede ampliar la dimensión de geografía (DIM_GEOGRAFIA) incorporando datos territoriales y geográficos, como longitudes y latitudes de departamentos y municipios. De manera similar, la dimensión de instituciones educativas (DIM_IIEE) podría extenderse con datos de geolocalización, brindando una interpretación gráfica de las incidencias de las variables analizadas. Lo mismo aplica a la dimensión de condiciones institucionales (DIM_COND_INST), que puede abarcar datos de condiciones estructurales y funcionales.

En la tabla de hechos relacionada con la estrategia EPA (FACT_EPA), que contiene la variable respuesta, se destaca su flexibilidad para analizar cualquier estrategia académica. Puede incluir referencias numéricas de valoración, como las pruebas ICFES SABER 11 o PRUEBAS SABER T&T. A un nivel más detallado, se pueden aplicar

valoraciones a nivel de institución educativa en relación con sus propias condiciones institucionales.

Respecto al análisis de datos desarrollado en esta monografía, existe un amplio margen para fortalecer la validación de las variables mediante métodos estadísticos más robustos. Específicamente, se pueden implementar técnicas de validación más complejas que integren datos históricos y modelos mejor entrenados, así como aplicar pruebas de normalidad, homogeneidad y correlación para asegurar la consistencia de los datos. Además, se propone avanzar hacia un enfoque analítico no solo centrado en la predicción, sino también en la evaluación profunda de la utilidad y pertinencia de cada variable dentro del modelo. Este tipo de análisis permitiría comprender con mayor precisión el peso real de cada dimensión en el desempeño educativo, y no únicamente su capacidad predictiva, abriendo así nuevas posibilidades para la formulación de políticas educativas basadas en evidencia.

Adicionalmente, es necesario reconocer las limitaciones presentadas durante el proceso de modelado, especialmente en lo relacionado con la disponibilidad y volumen de datos. La baja cantidad de registros impidió la aplicación de modelos más complejos que requieren un mayor número de observaciones para evitar el sobreajuste y garantizar resultados estadísticamente representativos. Esta limitación restringió el análisis principalmente a un modelo de regresión lineal múltiple. Sin embargo, en trabajos futuros se propone explorar y comparar otros enfoques de modelado, como árboles de decisión, bosques aleatorios, redes neuronales o modelos de ensamble como Gradient Boosting, que podrían capturar relaciones no lineales o interacciones complejas entre variables. La aplicación de estos modelos, acompañada de técnicas como la validación cruzada, permitiría obtener estimaciones más robustas y ofrecer una mayor comprensión del fenómeno educativo desde una perspectiva predictiva y explicativa.

12. Referencias

- A. Sorour, A. S. Atkins, C. F. Stanier and F. D. Alharbi. (2020). The Role of Business Intelligence and Analytics in Higher Education Quality: A Proposed Architecture- 2019 International Conference on Advances in the Emerging Computing Technologies (AECT), Al Madinah Al Munawwarah, Saudi Arabia, 2020, pp. 1-6, DOI: 10.1109/AECT47998.2020.9194157
- Agyei, E. A., Annim, S. K., Acquah, B. Y. S., Sebu, J., & Agyei, S. K. (2024). Education infrastructure inequality and academic performance in Ghana. *Heliyon*, 10(14).
- Alcaldía de Montería. (2021, marzo 1). Montería, la mejor en la Región Caribe en Pruebas Saber 11. Alcaldía de Montería. <https://www.monteria.gov.co/publicaciones/2686/monteria-la-mejor-en-la-region-caribe-en-pruebas-saber-11/>
- Alcaldía de Montería. (2022, noviembre 30). Montería avanza en Calidad Educativa. Colegios oficiales se destacan en Saber 11. Alcaldía de Montería. <https://www.monteria.gov.co/publicaciones/3771/monteria-avanza-en-calidad-educativa-colegios-oficiales-se-destacan-en-saber-11/>
- Araiza Lozano, Miguel Ángel. (2021). Factores socioeconómicos asociados al rendimiento académico de estudiantes universitarios. *Dilemas contemporáneos: educación, política y valores*, 9(1), 00007. Epub 03 de noviembre de 2021. <https://doi.org/10.46377/dilemas.v9i1.2831>
- Borja Naranjo, G. M., Martínez Benítez, J. E., Barreno Freire, S. N., & Haro Jácome, O. F. (2021). 3.- Factores asociados al rendimiento académico: Un estudio de caso. *Revista EDUCARE - UPEL-IPB - Segunda Nueva Etapa 2.0*, 25(3), 54–77. <https://doi.org/10.46498/reduipb.v25i3.1509>
- C. F. Olivos Seclen, “Inteligencia de negocios en el sector educativo de nivel superior,” *Bachiller*, Facultad de Ingeniería, Universidad Católica Santo Toribio de Mogrovejo, Chiclayo, Perú, 2023. [En línea]. Disponible en: <http://hdl.handle.net/20.500.12423/6219>

Cabrera, M. A. M. (2016). Determinantes que influyen en el rendimiento académico: un estudio aplicado para Colombia a partir de las Pruebas Icfes - estudio aplicado para Colombia a partir de las Pruebas Icfes - Saber 11 Saber 11. [Universidad de La Salle Universidad de La Salle].

<https://ciencia.lasalle.edu.co/cgi/viewcontent.cgi?article=1231&context=economia>

Caicedo, C. O. (2022). Desarrollo de una guía metodológica como herramienta de apoyo en la elaboración de una arquitectura de inteligencia de negocios con relación al área de salud del régimen subsidiado y contributivo en la Asociación Indígena del Cauca-EPS.

CORPORACIÓN UNIVERSITARIA AUTÓNOMA DE CAUCA.

Cárdenas, C. S. S., & León, J. H. Z. (2023). Uso del Big Data y la Analítica de Datos como Instrumentos Facilitadores en la Toma de Decisiones para las Empresas [Universidad Santo Tomás].

<https://repository.usta.edu.co/bitstream/handle/11634/48600/Uso%20del%20Big%20Data%20y%20la%20Analítica%20de%20Datos%20como%20Instrumentos%20Facilitadores%20en%20la%20Toma%20de%20Decisiones%20para%20las%20Empresas..pdf?sequence=1&isAllowed=y>

Chicaiza, M., Patricio, R., Armas, Q., & Elizabeth, F. (2022). Propuesta tecnológica de herramientas de business intelligence para la mejora de la gestión académica en la educación superior [Pontificia Universidad Católica del Ecuador].

<https://repositorio.pucesa.edu.ec/handle/123456789/3733>

Cuellar, J. P. V. (2022, marzo 25). Icfes pone a disposición una Caja de Herramientas para que estudiantes practiquen. Diario La República. <https://www.larepublica.co/economia/icfes-pone-a-disposicion-una-caja-de-herramientas-para-que-estudiantes-practiquen-3330584>

De-La-Hoz, Enrique J., De-La-Hoz, Efraín J., & Fontalvo, Tomás J.. (2019). Methodology of Machine Learning for the classification and Prediction of users in Virtual Education

Environments. Información tecnológica, 30(1), 247-254. <https://dx.doi.org/10.4067/S0718-07642019000100247>

Díaz, J. E. C., & Quiroga, C. C. D. (2022). ProtoData: Prototipo para la Generación de Estadísticas y Tableros de Control de Datos Abiertos Gratuita [Universidad Piloto de Colombia]. <http://repository.unipiloto.edu.co/handle/20.500.12277/12134>

Díaz, J. E. C., & Quiroga, C. C. D. (2022). ProtoData: Prototipo para la Generación de Estadísticas y Tableros de Control de Datos Abiertos Gratuita [Universidad Piloto de Colombia]. <http://repository.unipiloto.edu.co/handle/20.500.12277/12134>

Duplaa, S. I. (2022). Diseño, implementación y evaluación MyFEPS de un modelo de Business Intelligence para empresas Pymes [Universidad de Belgrano - Facultad de Ingeniería y Tecnología Informática - Licenciatura en Sistemas de Información]. <http://repositorio.ub.edu.ar/handle/123456789/10349>

García-Herrero, M., Rodríguez-Conde, M.-J., & Martínez-Abad, F. (2024). Factores de calidad docente asociados a la equidad educativa: formación del profesorado y estrategias docentes. *Revista Electrónica Interuniversitaria de Formación del Profesorado*, 27(1), 75–88. <https://doi.org/10.6018/reifop.595181>

Gómez-Duque, L. Á., Daza-Torres, J. D., & Arias-Pérez, J. (2023). Inteligencia de negocios y agilidad organizacional: ¿Son relevantes la toma de decisiones racional e intuitiva? *Estudios gerenciales*, 181–191. <https://doi.org/10.18046/j.estger.2023.167.5542>

Gonzaga, B. R. A. (2021). Inteligencia de negocios para la toma de decisiones: Un enfoque desde la dirección estratégica de instituciones educativas. *Revista Scientific*, 6(19), 295-312.

Gonzales, J. A. (2021). DISEÑO Y METODOLOGÍA DE LA INVESTIGACIÓN. ENFOQUES CONSULTING EIRL. https://gc.scalahed.com/recursos/files/r161r/w26022w/Arias_S2.pdf

- Gutiérrez, Y. Á. T., Ramos, M. C. M., & Morinson, M. L. V. (2022). Análisis multivariado de los factores socio-económicos asociados al rendimiento en las pruebas saber pro 2016: El caso de los estudiantes de Licenciatura en Matemáticas en Colombia: Multivariate analysis of the socioeconomic factors associated to the performance in the pro saber 2016 tests: The case of the students of the Bachelor of Mathematics in Colombia. *South Florida Journal of Development*, 3(1), 928–939. <https://doi.org/10.46932/sfjdv3n1-071>
- Hanushek, E., & Woessmann, L. (2021). Education and Economic Growth. *Oxford Research Encyclopedia of Economics and Finance*.
<https://doi.org/10.1093/acrefore/9780190625979.013.651>
- Harby, A. A., & Zulkernine, F. (2022). From data warehouse to lakehouse: A comparative review. *2022 IEEE International Conference on Big Data (Big Data)*.
- Hernandez Sampieri, R. (2006). *Metodologia de La Investigacion*. McGraw-Hill Companies.
- Herrera, J. V. E., Goyez, J. P. L., & Zabala, W. (2022). Business Intelligence & Data Analytics aplicado al proceso de seguimiento curricular en la universidad UPEC. *Minerva Journal*.
- Hinojosa, M., Derpich, I., Alfaro, M., Ruete, D., Caroca, A., & Gatica, G. (2022). Procedimento de agrupamento de alunos de acordo com o risco de evasão para melhorar a gestão estudantil no ensino superior. *Texto Livre*, 15, e37275. <https://doi.org/10.35699/1983-3652.2022.37275>
- IBM. (2024). ¿Qué es el análisis de datos exploratorio? IBM. <https://www.ibm.com/es-es/topics/exploratory-data-analysis>
- Icfes. (2023). “Evaluar para avanzar” 3-11. [icfes.gov.co](https://evaluarparaavanzar311.icfes.gov.co/).
<https://evaluarparaavanzar311.icfes.gov.co/>
- ICFES. (2023). Guia de preguntas frecuentes.
<https://www.icfes.gov.co/documents/39286/2163563/Guía+de+Preguntas+Frecuentes+2023.pdf>

Icfes. (2023a). “Evaluar para avanzar” 3° a 11°. icfes.gov.co.

<https://www.icfes.gov.co/evaluarparaavanzar-3-11>

ICFES. (2023a). Infografía general 2023. icfes.gov.co.

<https://www.icfes.gov.co/documents/39286/23440432/infografía+general2023ok.pdf/9f1731c4-5335-0e92-d4f5-8b50ac204a11?version=1.0&t=1680540707548>

Inmon, W. H. (2013). Building the Data Warehouse. Lulu.com.

Junior, C. G. M., Cavalcante, T. S. B., Vieira, A. K., & Vieira, R. K. (2024). Application of Business Intelligence in the Brazilian Educational System. *The Journal of Engineering and Exact Sciences*, 10(7), 20045-20045.

Khatibi, V., Keramati, A., & Shirazi, F. (2020). Deployment of a business intelligence model to evaluate Iranian national higher education. *Social Sciences & Humanities Open*, 2(1), 100056. <https://doi.org/10.1016/j.ssaho.2020.100056>

Kimball, R., & Ross, M. (2013). *The data warehouse toolkit: The definitive guide to dimensional modeling* (3a ed.). John Wiley & Sons.

Laboratorio de Economía de la Educación (LEE) de la Pontificia Universidad Javeriana. (2022). Informe No. 64. Primeros resultados de ““Evaluar para avanzar””: la nueva estrategia del ICFES. Disponible en <https://lee.javeriana.edu.co/documents/5581483/5629089/Informe64-Evaluar-para-Avanzar-2022.pdf>

Liu, J., et al. (2022). Socioeconomic status and academic achievement in primary and secondary education: A meta-analytic review. *Educational Psychology Review*, 34(4), 2867-2896.

López, A., & Andrés, P. (2023). Business intelligence model powered by data science applied to enterprise resource planning [Universidad de Investigación de Tecnología Experimental Yachay]. <http://repositorio.yachaytech.edu.ec/handle/123456789/656>

- Luque Salluco, V. (2021). Modelo de Inteligencia de Negocio en la Administración Pública. INF-FCPN-PGI Revista PGI, (8), 201–205. Recuperado a partir de https://ojs.umsa.bo/ojs/index.php/inf_fcpn_pgi/article/view/85
- Marciniak, R., & Cáliz Rivera, C. (2021). A system of indicators for the quality assessment of didactic materials in online education. *International Review of Research in Open and Distributed Learning*, 22(1), 180-198.
- Medina, M. A. A. (2023). IMPLEMENTACIÓN DE UNA SOLUCIÓN BUSINESS INTELLIGENCE PARA LA MEJORA DE LA TOMA DE DECISIONES EN MODERNIZACIÓN DE LA GESTIÓN DE LOS RECURSOS HÍDRICOS, AÑO 2022 [UNIVERSIDAD PERUANA DE LAS AMÉRICAS]. <http://repositorio.ulasamericas.edu.pe/bitstream/handle/123456789/3730/1%20TESIS%20ATAULLUCO%20MEDINA.pdf?sequence=1&isAllowed=y>
- Mendoza-Rivera, R. D. (2022). Inteligencia de Negocios para Agilizar la Toma de Decisiones en la Gestión de Pacientes de Policlínicos de Salud. Memorias de la Vigésima Primera Conferencia Iberoamericana en Sistemas, Cibernética e Informática: CISCI 2022.
- Ministerio de Educación Nacional. (2019). Decreto 1330 de julio 25 de 2019. Por el cual se sustituye el Capítulo 2 y se suprime el Capítulo 7 del Título 3 de la Parte 5 del Libro 2 del Decreto 1075 de 2015 - Único Reglamentario del Sector Educación. Recuperado de https://www.mineducacion.gov.co/1780/articles-387348_archivo_pdf.pdf
- Ministerio de Tecnologías de la Información y las Comunicaciones. (2023). Datos abiertos. mintic. <https://gobiernodigital.mintic.gov.co/portal/Iniciativas/Datos-abiertos/>
- Ong, I., Siew, P., & Wong, S. (2011). A Five-Layered Business Intelligence Architecture. *Communications of the IBIMA*, 1–11. <https://doi.org/10.5171/2011.695619>
- Ortigoza, Analía, Prchal, Abigail, & Canova-Barrios, Carlos. (2024). Academic trajectories of nursing students at the National University of Tucumán through the analysis of academic

performance indicators. *Enfermería Global*, 23(74), 390-409. Epub 30 de julio de 2024. <https://dx.doi.org/10.6018/eglobal.600331>

Osorio-Sanabria, M.-A., Amaya-Fernández, F. O., & González-Zabala, M. P. (2021). Políticas, normas y estrategias que fomentan los datos abiertos en Colombia: un análisis de literatura. *Revista virtual Universidad Católica del Norte*, 62, 155–188. <https://doi.org/10.35575/rvucn.n62a7>

outsourcando. (2013). Las 5 Clases de Usuario que tiene toda Solución de Business Intelligence (BI). [outsourcando.blogspot.com](https://outsourcando.blogspot.com/2013/04/cinco.clases.usuario.BI.html). <https://outsourcando.blogspot.com/2013/04/cinco.clases.usuario.BI.html>

Rodríguez Garcés, Carlos, Espinosa Valenzuela, Denisse, Padilla Fuentes, Geraldo, & Suazo Ruíz, Claudia. (2022). Entre el talento académico y la segmentación socioeducativa: admisión universitaria de estudiantes Top 10% Ranking en Chile. *Revista Cubana de Educación Superior*, 41(2), . Epub 10 de mayo de 2022. Recuperado en 15 de julio de 2025, de http://scielo.sld.cu/scielo.php?script=sci_arttext&pid=S0257-43142022000200004&lng=es&tlng=pt.

Rodríguez Rosero, D. D., Ordoñez Ortega, R. E., & Hidalgo Villota, M. E. (2021). Determinantes del rendimiento académico de la educación media en el departamento de Nariño, Colombia. *Lecturas de Economía*, (94), 87-126.

Rodríguez Vargas, M. M., Regalado, R., & Imamura Díaz, J. I. (2023). La evaluación del desempeño a los estudiantes de electricidad en la práctica preprofesional. *Horizonte pedagógico (La Habana)*, 12(1), 66-77.

Romero, M. I. L. (2023). Diseño de una arquitectura de inteligencia de negocios para apoyar la toma de decisiones en el sector público de Colombia: caso de estudio Antioquia [Universidad EAN].

<https://repository.universidadean.edu.co/bitstream/handle/10882/12863/LopezMaria2023.pdf?sequence=1&isAllowed=y>

Ruiz-Quintero, I. (2014). “Inteligencia de Negocios al Proceso de La Evaluación Docente”. In *Vestigium Ire*. Vol. 8, pp. 206-214.

Sánchez Bracho, M., Fernández, M., & Díaz, J. (2021). Técnicas e instrumentos de recolección de información: análisis y procesamiento realizado por el investigador cualitativo. *Revista Científica UISRAEL*, 8(1), 107–121. <https://doi.org/10.35290/rcui.v8n1.2021.400>

Seclen, O., & Fabricio, C. (2023). Inteligencia de negocios en el sector educativo de nivel superior [Universidad Católica Santo Toribio de Mogrovejo]. <https://tesis.usat.edu.pe/handle/20.500.12423/6219>

Sedkaoui, S. (2018). *Data Analytics and Big Data*. Standards Information Network.

Sharda, R., Turban, E., Delen, D.(2018). *Business intelligence, analytics, and data science*. Pearson Educación. <https://www-ebooks7-24-com.bdbiblioteca.universidadean.edu.co/?il=6301>

Simón Medina, N., Abellán López, M. Á., & Cisneros Britto, J. C. (2022). *Pedagogía Terapéutica y uso de las TIC: un análisis de segmentación en Castilla-La Mancha*.

Solano Mesa, J. E., & Orjuela Herrera, J. F. (2022). *Tablero de control para el análisis sobre la efectividad de los cultivos y su relación con el clima a partir de datos abiertos en Colombia* [Universidad Antonio Nariño]. <http://repositorio.uan.edu.co/handle/123456789/7389>

Solano Mesa, J. E., & Orjuela Herrera, J. F. (2022). *Tablero de control para el análisis sobre la efectividad de los cultivos y su relación con el clima a partir de datos abiertos en Colombia* [Universidad Antonio Nariño]. <http://repositorio.uan.edu.co/handle/123456789/7389>

Valencia, J. A. L. (2022). Aplicación de Inteligencia de negocios para la toma de decisiones en la Institución Educativa Manuel María Mosquera del municipio de Puracé – Cauca [Fundación Universitaria de Popayán - FUP].
<https://unividadafup.edu.co/repositorio/files/original/ed41c80fdb671076691663115b9c0b52.pdf>

Valiente, A. J. A., de González, M. G. O., & Zepeda, M. L. (2014). Incidencia del factor socioeconómico en el rendimiento académico de los/as estudiantes de la Carrera de Licenciatura en Ciencias de la Educación de la Universidad de El Salvador [UNIVERSIDAD DE EL SALVADOR].
<https://ri.ues.edu.sv/id/eprint/6381/1/Incidencia%20del%20factor%20socioeconómico%20en%20el%20rendimiento%20académico%20de%20losas%20estudiantes%20de%20la%20Carrera%20de%20Licenciatura%20en%20Ciencias%20de%20la%20Educación%20de%20la%20Universidad%20de%20El%20Salvador.pdf>

Valladares Fuente, Fernando Emilio, & Posso Pacheco, Richar Jacobo. (2022). Indicadores de la educación ambiental para el profesional del deporte. *Ciencia y Deporte*, 7(1), 45-59. Epub 03 de abril de 2022. <https://dx.doi.org/10.34982/2223.1773.2022.v7.no1.004>

Vargas Agudelo, F. A., Giraldo Mejía, J. C., Delgado González, I. A., & Parra Valencia, J. A. . (2022). Modelo de inteligencia de negocios para evaluar el impacto de la investigación y la innovación en programas de ingeniería en Colombia. *Encuentro Internacional De Educación En Ingeniería*. <https://doi.org/10.26507/paper.2343>

Velosa, A., & Pabon, G. (2021). *Business Intelligence and Its Big Evolution*. Preprints.
<https://www.preprints.org/manuscript/202103.0584/v1>

Vera Menendez, L. V. (2022). Indicadores de influencia en el rendimiento escolar postpandemia. *Social Innova Sciences*, 3(3), 30-41. <https://doi.org/10.58720/sis.v3i3.98>

Vicente, E. H. J., Pablo, L. G. J., & Andrés, Z. V. W. (2022). Business Intelligence & Data

Analytics aplicado al proceso de seguimiento. *Minerva Journal*, 3, 10.

<https://scholar.archive.org/work/nevm2lewx5cv5e64f7z2pslnq4/access/wayback/https://minerva.autanabooks.com/index.php/Minerva/article/download/75/233>

Vidales, P., & Cleivert, W. (2022). Business Intelligence para el proceso de toma de decisiones

en gestión del servicio de Asociación Educativa Global, Colombia S.A.S 2022

[Universidad César Vallejo]. <https://repositorio.ucv.edu.pe/handle/20.500.12692/109448>

Villegas-Ch, W., Palacios-Pacheco, X., & Luján-Mora, S. (2020). A business intelligence

framework for analyzing educational data. *Sustainability*, 12(14), 5745.